Vol. 28, No. 5 October, 1944

Journal of Applied Psychology

EDITED BY: DONALD G. PATERSON, UNIVERSITY OF MINNESOTA

Consulting Editors

PAUL S. ACHILLES, *Psychological Corporation;* WALTER V. BINGHAM, *A.G.O., War Department;* HAROLD E. BURTT, *Ohio State University;* ARTHUR I. GATES, *T. C. Columbia University;* JOHN G. JENKINS, *University of Maryland;* IRVING LORGE, *T. C. Columbia University;* QUINN MCNEMAR, *Stanford University;* WILLARD C. OLSON, *University of Michigan;* JAMES P. PORTER, *Ohio University;* EDWARD K. STRONG, JR., *Stanford University;* MORRIS S. VITELES, *University of Pennsylvania;* JOSEPH ZUBIN, *N. Y. Psychiatric Institute.*

Table of Contents

Tenth Nation-wide Social Experimental Survey: H. C. LINK 363

An Engineering and Physical Science Aptitude Test: C. H. GRIFFIN AND H. BOROW 376

A Study of the Reliability of the Minnesota Rate of Manipulation Test by the Split-half and Test-retest Methods: J. TUCKMAN 388

A Preliminary Report on Some New Tests of Musical Ability: R. W. LUNDIN 393

Leadership as Related to the Bernreuter Personality Measures: IV. An Item Analysis of Responses of Adult Leaders and Non-Leaders: N. G. HANAWALT AND H. M. RICHARDSON 397

Simplifying the Scoring Technique of the Bernreuter Personality Inventory: H. KEMPFER 412

Simplified Scoring of the Bernreuter Personality Inventory: D. C. MCCLELLAND 414

An Empirical Evaluation of Study Habits in Elementary German: J. R. WITTENBORN AND R. P. LARSEN 420

The Use of the Free Association Technique in the Investigation of the Stimulus Value of Trade Names: J. P. FOLEY, JR. 431

Job Specifications in Applied Psychology 436

Book Reviews 438

Published Bi-monthly by The American Psychological Association, Inc.
With the Cooperation of The American Association for Applied Psychology
Prince and Lemon Sts., Lancaster, Pa., and Northwestern University, Evanston, Illinois

Entered as second-class matter, August 19, 1943, at the post office at Lancaster, Pa., under the Act of March 3, 1879
Copyright, 1944, by The American Psychological Association, Inc.

Journal of Applied Psychology

Vol. 28, No. 5 October, 1944

The Tenth Nation-wide Social Experimental Survey

Henry C. Link

The Psychological Corporation, New York City

In keeping with the professional standards of psychology, all surveys by The Psychological Corporation, since their beginning in 1932, have included in their reports these critical facts:

1. *Date of Interviews.* A perfect survey, assuming this were possible, gives answers which are true only as of the date of the interviews. The interviews in the present survey were made April 10th to April 24th.

2. *Number of Interviews.* Unless the number of interviews is given, the probable statistical error is unknown. In this study 5000 interviews were made with an equal number of men and women, with two questionnaires, each used with half the sample.

3. *Place and Manner.* The Psychological Corporation has always maintained that no accurate cross section of the public could be obtained by mail or telephone interviews. Therefore, all interviews are made personally, with men and women in their homes, but only one in any one home. Homes are assigned by streets, blocks, and sections. Interviewing people at public places, offices, factories, stores, et cetera, is done only when a special, rather than a comprehensive, sample is desired.

4. *Localities and Sections.* The distribution by cities, towns, or rural localities is given either in detail or summary. The present study was made in 118 cities and towns according to the distribution given hereafter.

5. *Number of Interviewers and Supervisors.* The present 5000 interviews were conducted by 412 interviewers and 121 psychologists acting as local supervisors. This means an average of only 12 interviews per interviewer. In order to assure the widest distribution of risks due to interviewer bias or incompetence, it has always been our practice to limit each interviewer to about 20 or 30 interviews. This also helps to prevent stereotyped results such as even the best interviewers tend toward after the first dozen or so interviews.

6. *Questions and Questionnaires.* Either the questionnaires themselves or the exact questions are always given. In the absence of these data, the results are not worth attention.

7. *Per Cents.* The total number of interviews on which the per cents are based is always made clear. All per cents, including don't knows or no answer, are based on the number of people interviewed.

8. *Socio-Economic Groups.* We have consistently used the decile system as used by the Brookings Institute. However, we have grouped these deciles as follows:

A Group—the highest 10% by income, education, occupation, and living conditions.

B Group—the next 30% by these criteria, largely the white collar, professional, and semi-professional class.
C Group—the next 40%, mostly factory, shop, skilled and semi-skilled workers, transportation, et cetera.
D Group—the last 20%, mostly unskilled, foreign language, or poorest tenement residents.

The uncertainties of public opinion, buying, and advertising surveys are magnified by the omission of any of the above facts.

Distribution of 5000 Interviews

Geographic Areas	#	%	Socio-Econ. Groups	#	%
East	1850	37	A	500	10
Mid-West	1574	31	B	1500	30
South	1030	21	C	2000	40
Far West	546	11	D	1000	20
Total	5000	100	Total	5000	100

City Size	#	%	Sex	#	%
500,000 and over	1530	31	Men	2500	50
100,000–500,000	1040	21	Women	2500	50
25,000–100,000	980	19	Total	5000	100
2,500–25,000	1450	29			
Total	5000	100			

Race and Class Prejudice

The subject of race and class prejudice is an unusually difficult one for investigation by personal interview. The questions were framed with the assistance of leaders in the field of racial relations. A total of forty different questions, or different wordings of questions, were tested in a series of seven tests ranging from twenty to fifty-five interviews each.

Religion of Respondents. As a check on our sampling of interviews, as well as to help in evaluating the results, it was considered desirable to find out the religious affiliation or background of all persons interviewed. Therefore, the following two questions were asked toward the end of the interview: (Questions are numbered for reference and do not necessarily indicate the order in the questionnaire.)

Q. 1. "By the way, would you object to telling me your religious denomination?"
Q. 2. "Are you active, or a church member?"

Answers	Total	East	Mid-West	South	Far West
Protestant	63%	51%	65%	80%	66%
Catholic	19	27	18	8	15
Jewish	9	17	7	4	3
Other	2	1	2	2	3
Refused	1	1	1	—	1
None	6	3	7	6	12
Total Interviews	2500	925	787	515	273

Answers	By Socio-Econ. Groups A	B	C	D
Protestant	68%	63%	62%	63%
Catholic	10	15	21	24
Jewish	13	13	9	3
Other	3	1	2	2
Refused	1	1	1	*
None	5	7	5	8
Total Interviews	250	750	1000	500

* Less than .5%.

It should be noted that only 1 per cent of the people refused to answer the question about their religion, and only 6 per cent said that they had none. In answer to the second question: "Are you active or a church member?", the results were as follows:

Answers	Active	Inactive
Protestant	61%	39%
Catholic	81	19
Jewish	43	57

This is one of the most interesting and significant of all results, though any interpretations must remain largely speculative.

Priming the Pump. First of all, two questions were asked to introduce the subject and to get people in a frame of mind suitable for the more important questions to follow. These questions were not intended to measure either the intensity of the prejudices involved or their extent. Such a measurement is quite beyond the scope of the usual public opinion survey, or even of an unusual one, for that matter. Therefore, the results should be taken with these qualifications in mind.

Q. 3. "As you know, there are different religions and races in the U. S. which some people are prejudiced against, the Protestants, the Catholics, the Jews, the Negroes, etc. Against which *one* of these groups, *if any*, do you think prejudice or feeling has *increased the most?*"

Answers	Total	East	Mid-West	South	Far West
Jews	37%	45%	37%	26%	35%
Negroes	31	25	32	40	30
Catholics	3	3	3	3	4
Protestants	*	*	*	1	—
None	15	14	14	12	20
Don't know	16	15	16	19	13

* Less than .5%.

By socio-economic groups there were no important differences comparable to the above differences by geographic sections. If we had the results of these questions from a previous study, we might have some indications of trends. As it is, the results are only the roughest kind of an indication.

Class Prejudice. The question in regard to class prejudice presented problems which no amount of testing could wholly eliminate. The category Fascists was omitted from this question because, obviously, we are at war with them.

Q. 4. "Also, there are certain social classes in the U. S. which some people are prejudiced against. Against which of the following, if any, do you think prejudice or feeling has increased the most: against Labor Unions; against Employers; against the Communists; against the Politicians?"

		By Socio-Econ. Groups			
Answers	Total	A	B	C	D
Labor Unions	48%	63%	57%	45%	35%
Employers	4	4	3	3	5
Communists	15	17	14	15	15
Politicians	20	19	20	21	20
None	5	4	5	5	5
Don't know	15	6	8	17	24

Socio-Economic Groups. Even though we regard the answer to this question as only a straw in the wind of public opinion, the great difference between the per cent naming Labor Unions and the per cent naming Employers, in all income groups, may have unusual significance at this time. The A and B groups are primarily executive, professional, semi-professional, office workers, and white-collar workers generally. The C and D groups, 60 per cent of the total, are primarily factory workers, skilled and unskilled labor, teamsters, etc. There were no significant differences between sections of the country, except that the South registered less prejudice against Politicians whereas the East registered greater prejudice against Communists.

Race Vs. Class Hatreds. Having registered in people's minds, by these two questions, the categories of *race* on the one hand and *class* on the other, the following question was asked.

Q. 5. "Which do you think is the more serious in the U. S., and the more dangerous to its future.

a. *Race hatreds:* that is, the hatred toward racial groups like the Negroes and the Jews; *or*

b. *Class hatreds:* that is, the hatreds between labor unions and employers, between Capitalists and Communists, the common man and the wealthy man;

c. *Or* do you think neither is dangerous; or both equally dangerous?"

		By Socio-Econ. Groups			
Answers	Total	A	B	C	D
Race hatreds	34%	35%	35%	34%	34%
Class hatreds	24	28	25	23	21
Both equally	25	26	27	25	20
Neither	7	6	6	8	7
Don't know	10	5	7	10	18

There are several remarkable points about these results. In the first place, 90 per cent of the people have a definite opinion, since only 10 per cent of them answered, don't know. This problem has probably never been put to them in this formal and specific manner before. From their answers it is evident that we are touching here on a dynamic situation. People have become conscious of class as well as of race conflict. Almost as many people consider class hatreds dangerous as consider race hatreds dangerous, 49 per cent as compared with 59 per cent. Only 7 per cent consider neither dangerous. Finally, there is substantial agreement between the opinions of the various socio-economic groups.

Race Riots. As some clue to the seriousness with which the American people regard this situation, we asked the following:

Q. 6. "Do you feel that race riots in the U. S., the molesting of churches and synagogues, and attacks on members of a different race, are going to happen more often or less often in the next few years?"

Answers	Total	East	Mid-West	South	Far West
More often	40%	34%	41%	50%	36%
Less often	34	40	32	29	32
Don't know	26	26	27	21	32

It is interesting to see that the South, and next the Mid-West, are the sections in which the expectation of further race riots is highest, even though the most serious riots in recent years were in the North and the Far West. By socio-economic groups, the expectancy of further riots is slightly higher in the A and B groups than in the C and D groups.

Remedies. Many questions and varieties of questions dealing with possible remedies were tested. Included in these test questions were many agencies such as the President's Fair Employment Practices Commission, the Inter-Racial Council of Churches, and others. The percentage of people who had heard of these organizations or who regarded their efforts as effective was very small. Hence, we were compelled to classify the remedies into broad categories, as follows:

Q. 7. "As a means of reducing race prejudice and class hatreds, which *one or two* of the following *could do* the most good? a. The grammar and high schools; b. The Churches and Sunday Schools; c. Public meetings and speeches; d. The labor unions and employers; e. Gov't laws and Gov't bureaus."

Answers	Total	By Socio-Econ. Groups A	B	C	D
Churches and Sunday Schools	55%	50%	52%	58%	57%
Grammar and High Schools	54	64	62	50	45
Gov't Laws and Gov't Bureaus	21	13	18	22	27
Public meetings and speeches	14	15	14	13	14
Labor unions and employers	12	19	13	11	11
Don't know	2	2	2	2	2

These per cents add up to more than 100 because some people mentioned more than one remedy. It is interesting to see that, although in the last year or two the most vocal activities against race prejudice have stemmed from Government Bureaus and Officials, and public meetings and speeches, mostly through the channels of specially organized pressure groups, these are considered the least effective remedies. Overwhelmingly, the basic channels of the common schools and the Churches and Sunday Schools are considered most important.

National or Local Remedies? The attempts to handle racial problems on a wholesale and national basis have aroused such antagonisms that this has become an important current political issue. The Town Hall Meeting of the Air, a few months ago, conducted an interesting discussion of just this question, namely, whether racial problems can be handled best by the Federal Government or by local Government and local authorities. We asked the following question in this survey:

Q. 8. "All in all, how can these problems be handled best? a. On a national basis, that is, by the Federal Government; b. On a local basis, by each state or city handling its own problems."

		By Socio-Econ. Groups			
Answers	Total	A	B	C	D
Local basis	59%	69%	62%	60%	48%
National basis	30	22	29	29	40
Both	2	3	2	2	1
Don't know	9	6	7	9	11

The decided majority thought that these problems could be dealt with better on a local than on a national basis, and by the State and local community than by the Federal Government. This was true of all socio-economic groups, except the D group, where opinions were more evenly divided. It should be noted that only 9 per cent of all people to whom this question was put did not have a positive opinion.

Wages and the Cost of Living

In view of the controversy over wages and the cost of living, the fight against inflation, the attempt to break the Little Steel Formula, and especially the conflicting claims as to which has gone up faster, wages or the cost of living, the beliefs of the people themselves should be interesting. Therefore, in one half of our sample we repeated a question which we have asked in three earlier studies, namely:

Q. 9. "Is your family more prosperous today than two years ago, less prosperous, or the same?"

Answers	In Oct. 1941	In Nov. 1942	In Oct. 1943	In Apr. 1944
More prosperous	38%	29%	29%	28%
About the same	47	47	46	48
Less prosperous	15	21	23	22
Don't know		3	2	2
Total Interviews	2000	2500	2500	2500

In spite of the increases in the cost of living, and in spite of the fact that this study came shortly after the heaviest tax payments in the history of the nation, 76 per cent said they were as prosperous or more prosperous than they were two years ago, or as many as in 1943 and 1942. Moreover, this increased or equal prosperity, while relatively greater in the C and D, or the large factory worker group, is almost as great in the B, or the large white-collar or office worker group. Only the A, or the wealthiest group, admit to a somewhat less prosperity and even there 64 per cent say they are as prosperous or more prosperous than they were. *In short, according to the statements of the people themselves, usually the husband or the wife, the great majority, regardless of class, have been as prosperous or more prosperous than two years earlier, throughout the war to date!*

		By Socio-Econ. Groups			
Answers in 1944	Total	A	B	C	D
More prosperous	28%	25%	27%	28%	32%
About the same	48	39	47	50	48
Less prosperous	22	32	24	20	18
Don't know	2	4	2	2	2

Some popular polls have asked questions on this subject, the wording of which was calculated to give a result quite different from those we have been obtaining. Therefore, we asked another question on this subject in the other half of the sample, another 2500 interviews, as follows:

Q. 10. "Are you and your family better off, or worse off, financially, than you were two or three years ago?"

		By Socio-Econ. Groups			
Answers	Total	A	B	C	D
Better off	37%	31%	31%	38%	48%
Neither	38	35	42	39	35
Worse off	23	31	26	21	17
Don't know	2	3	1	2	*

The answers here indicate even more strongly than in the other question that people are better off, financially, than they were, 37 per cent as compared with 28 per cent in the other question. The total who say they are better off or as well off is about the same, 75 per cent as compared with 76 per cent. The differences by socio-economic groups show a little more clearly a proportionately greater increase in prosperity as we go down the line. However, again the white-collar, A and B groups, about whose failure to participate in the many wage increases awarded by the W. L. B. there has been so much publicity, do not seem to feel so badly about their financial position.

Free Enterprise Vs. Government Control

In the October 1943 study we had a half dozen questions dealing with free enterprise vs. Government control. In the present study we repeated, for the sake of the trend, two of these questions.

Q. 11. "Who do you think can do the best job in straightening things out after the war: the government in Washington; business leaders; labor union leaders; or others?"

Answers	In Oct. 1941	In Nov. 1942	In Oct. 1943	In Apr. 1944
Government in Washington	47%	46%	42%	41%
Business leaders	26	25	28	23
Labor union leaders	5	8	8	6
All three together	7	6	9	13
Others or had no opinion	17	17	17	20
Total Interviews	2000	2500	2500	2500

The above per cents add up to more than 100 because some people named more than one agency. Reliance on Government as contrasted with business and labor union leaders is high, but it should be remembered that the question deals with a situation, namely, the war, in which the Government has naturally been compelled to play a tremendous role. The results by socio-economic groups are significant.

	By Socio-Econ. Groups			
Answers	A	B	C	D
Government in Washington	27%	41%	43%	45%
Business leaders	39	27	21	14
Labor union leaders	3	3	7	10
All three together	21	16	12	7
Others	8	7	6	5
Don't know	6	9	14	20
Total Interviews	250	750	1000	500

Government Management Vs. Private Management of Industry. One of the best single questions we have found as a measure on this subject, and which was asked in the October 1943 study and also in some of the studies made privately for clients, was this:

Q. 12. "If all manufacturing companies were run entirely by the Government, would you get more for your dollar than you do under the present system, or not so much, or about the same?"

		By Socio-Econ. Groups			
Answers	Total	A	B	C	D
Not so much	40%	62%	46%	35%	26%
More for the dollar	22	12	20	25	25
Same	17	11	15	19	22
Don't know	21	15	19	21	27

The results are almost exactly as they were in the October 1943 study. However, in that study we asked this question of a thousand college students also, and here the results were almost the reverse, 39 per cent saying that they would get more, 25 per cent not so much, and 26 per cent about the same. We intend to repeat this question in future studies.

1944 Election Issues

From a considerable number of issues tried out in our preliminary tests, eight were selected and the respondent was asked in each case whether or not he would like to have this issue discussed. The form of the question and the results were as follows:

Q. 13. "In the coming Presidential campaign which of the following subjects would you *ESPECIALLY* like the candidates to talk about? (Name each one below, ask 'Yes or No' and check.)

Percentage Answering Yes to Each Issue

Jobs after the war	86%	Government taxes	71%
The Peace Settlement	84	Regulating the unions	67
The cost of living	79	Race and religious prejudice	47
Cutting out unnecessary Gov't Bureaus	78	How to run the war	35

Jobs after the War. Jobs after the war continues, as in past studies, to be the one most important issue. The peace settlement is a close second.

Internationalism Vs. Isolationism. In February 1943 we made a survey using a depth interview to find out the extent to which people favored international cooperation or an isolationist policy. This study (an Experiment in Depth Interviewing on the issue of Internationalism Vs. Isolationism—Henry C. Link, *Public Opinion Quarterly*, summer issue 1943) showed that the great majority of the American people were in favor of international cooperation in spite of the heavy costs involved—costs which we took care to point out to them in the course of the interview. In the present study we repeated one of the questions asked in the study of February 1943, namely:

Q. 14. "You remember that, after the last war, the U. S. got out of Europe as quickly as possible, and stayed out. This time, in order to keep the peace, do you think the U. S. A. should take its full share in world affairs?"

Answers	Feb. '43	Apr. '44
Yes	80%	84%
No	7	7
Don't know	13	9

Most interesting is the fact that, among the rank and file of the people of America, there is no isolationist block or section. All sections, the Mid-West, the South, the East, the Far West, are equally and overwhelmingly committed to international cooperation. The results by sections of the country were as follows:

Answers	East	Mid-West	South	Far West
Yes	84%	82%	86%	82%
No	6	9	5	8
Don't know	10	9	9	10
Total Interviews	925	787	515	273

Pay As You Go Taxes. A hot issue sometime ago was whether or not Congress should levy 10 billion dollars more in taxes, and especially how such taxes should be levied. We asked the following two questions on this subject:

Q. 15. "It has been said that we need at least *10 billion dollars* a year *more* in taxes to help pay for the war. If these higher taxes are passed by Congress, how would you prefer to pay for them: (a) by larger deductions from your (or your husband's) wages; or (b) by a 10% sales tax on everything you buy except a few of the foods?"

		By Socio-Econ. Groups			
Answers	Total	A	B	C	D
Wage Deductions	31%	29%	32%	31%	30%
10% Sales taxes	52	58	56	51	47
Don't know	17	13	12	18	23
Total Interviews	2500	250	750	1000	500

These results confirm the results of polls we conducted in June 1941 and June 1942, among 5000 families in each case, when the American public, regardless of income group, overwhelmingly preferred a sales tax to any other form of taxation. Our findings in the past have special significance because other polls, until recently, have obtained radically different results. The form of question used has everything to do with the kind of answer obtained. Following the above question we asked this question:

Q. 16. "Do you think that Congress should or should not pass a bill for 10 billion dollars more?"

		By Socio-Econ. Groups			
Answers	Total	A	B	C	D
Should pass bill	44%	50%	48%	42%	38%
Should not pass bill	26	24	23	27	27
Don't know	30	26	29	31	35

While a large proportion of the people in all income groups are in favor of additional taxes, according to this question, the percentage is not as large as that found in some of the other popular polls.

This is probably due to the fact that our question was preceded by a question as to *how* the individual is going to pay these taxes. Obviously, not only the wording of the question is important to the answers obtained; the nature of the question preceding it and the ideas put in people's minds before the question is asked also have an important influence on the answers.

The Power of Suggestion. We thought we would make a little test of the power of suggestion in the form of two differently worded questions on the same subject. One of these questions was asked of half the sample, or 2500 men and women; the other one was asked of the other half of our sample, another 2500 men and women. The questions were:

Q. A. "You probably heard the saying, 'Don't change horses in midstream.' Do you think it would be a good thing, or a bad thing, to elect a *new* President next November?"

Q. B. "You have probably heard the saying, 'A new broom sweeps clean.' Do you think it would be a good thing, or a bad thing, to elect a *new* President next November?"

Answers	Q. A	Q. B
A good thing to change	38%	40%
A bad thing to change	45	45
Don't know	17	15

In considering these results it should be remembered that our sample is a city and not a small town and farm sample. However, it should be noted that the suggestion intended in these two questions did not displace the central conviction which people have about the old or a new President. The results by socio-economic groups, however, are also interesting.

Answers	A	B	C	D	Men	Women
A good thing to change	56%	45%	37%	29%	44%	36%
A bad thing to change	30	42	48	53	43	47
Don't know	14	13	15	18	13	17
Total Interviews	250	750	1000	500	1250	1250

As in all other polls, the lower the socio-economic status, the stronger the present incumbent's appeal. Socio-economic status, as used here, includes income, education, and other criteria. Lest this be misinterpreted it should be pointed out that students and teachers are usually found to be decidedly in favor of the present President. Interesting and significant for the election is the finding that women are much less in favor of changing than are the men.

Q. 17. "Before a candidate is nominated for President by either the Democrats or the Republicans, do you believe he should tell us what he thinks about subjects like the ones we just talked about; or do you think that he should wait until after he has been nominated?"

Answers	Total
Should tell us	78%
Wait after nomination	13
Don't know	9

The surprising thing here is that 13 per cent of the people say that the candidate should wait until after nomination to discuss issues.

Post War Pessimism

After two years of optimism in post-war conditions, the trend seems to have been definitely reversed. Following are a series of questions identical to those asked in 1941, 1942, and 1943, and their results:

Q. 18. "When this war is over do you think that the people of this country will be better or worse off than they are now?"

Answers	In Oct. 1941	In Nov. 1942	In Oct. 1943	In Apr. 1944
They would be better off	13%	40%	48%	31%
They would be worse off	69	37	32	47
Don't know	18	23	20	22
Total Interviews	2000	2500	2500	2500

Q. 19. "How about jobs; will there be more, fewer, about the same?"

Answers	In Oct. 1941	In Nov. 1942	In Oct. 1943	In Apr. 1944
There would be more jobs	8%	20%	26%	22%
There would be fewer jobs	74	50	46	51
About the same	11	18	20	17
Don't know	7	12	8	10

Q. 20. "Will taxes be higher, about the same, or lower?"

Answers	In Oct. 1941	In Nov. 1942	In Oct. 1943	In Apr. 1944
Taxes would be higher	67%	53%	48%	44%
Taxes would be about same	17	25	31	31
Taxes would be lower	9	11	13	14
Don't know	7	11	8	11

Q. 21. "Will wages be higher, about the same, or lower?"

Answers	In Oct. 1941	In Nov. 1942	In Oct. 1943	In Apr. 1944
Wages would be higher	10%	7%	8%	6%
Wages would be about same	20	24	26	26
Wages would be lower	60	59	60	60
Don't know	10	10	6	8

Q. 22. "Will our government be less democratic, more democratic, or the same as now?"

Answers	In Oct. 1941	In Nov. 1942	In Oct. 1943	In Apr. 1944
Gov't would be less democratic	26%	17%	19%	17%
Gov't would be more democratic	19	32	30	27
Gov't would be same as now	33	31	34	31
Don't know	22	20	17	25

Received July 21, 1944.

References*

1. P. G. Corby, S. Roslow, and W. H. Wulfeck. Consumer and opinion research: Experimental studies on the form of question. *J. appl. Psychol.*, 1940, **24**, 334–346.
2. Link, H. C. How many interviews are necessary for results of a certain accuracy? *J. appl. Psychol.*, 1937, **21**, 1–17.
3. The Psychological Corporation. A study of public relations and social attitudes. *J. appl. Psychol.*, 1937, **21**, 589–602.
4. Link, H. C. A study of public opinion and morale. *J. appl. Psychol.*, 1941, **25**, 636–645.
5. Link, H. C., and Freiberg, A. D. The problem of validity vs. reliability in public opinion polls. *Publ. Opin. Quart.*, 1942, **5**, 87–98.
6. Link, H. C. The eighth nation-wide social experimental survey. *J. appl. Psychol.*, 1943, **27**, 1–11.
7. Link, H. C. An experiment in depth interviewing on the issue of internationalism vs. isolationism. *Publ. Opin. Quart.*, 1943, **6**, 267–279.
8. Link, H. C. The ninth nation-wide social experimental survey. *J. appl. Psychol.*, 1944, **28**, 1–15.

* A complete set of references to studies published by The Psychological Corporation is available on request.

An Engineering and Physical Science Aptitude Test

Charles H. Griffin and Henry Borow

The Pennsylvania State College

The current national crisis has seen the training of great numbers of technical specialists whose services are urgently needed for the successful pursuit of the war. A significant share of the civil responsibility for training scientific personnel on an advanced educational level has been assigned to colleges and universities under the federal program known as Engineering, Science and Management Defense Training.[1] With the assistance of funds administered by the U. S. Office of Education, those institutions designated as training centers under ESMWT have offered tuition-free courses to qualified men and women in a large number of fields related to the production of war materials.

At The Pennsylvania State College, where a most extensive ESMWT program is being conducted, aptitude testing has been a standard part of the training procedures. The development and application of psychological tests are here carried out by the Personnel Service Division, originally a branch of ESMWT locally instituted to enhance the effectiveness of the training program by psychological research methods. One of the first duties with which the Personnel Service Division was charged was the development of a test battery to appraise aptitude for training in work of an engineering or physical science nature. Whereas many tests were available at the time, each sampling a relatively restricted area of aptitude for training in the physical sciences, no omnibus measure had as yet appeared designed specifically to assess a number of the abilities important to technical performance at the college level. The research of the armed services in this field will, in all probability, not be publicized until the close of the war.

Development of the Test

The Engineering and Physical Science Aptitude Test represents a compilation from a series of previously developed tests. The selection of the test parts was based on a physics criterion while validation of the final battery was determined with reference to attainment in a series of technical subjects. Male high school graduates enrolled in the Introductory Engineering Subjects course of The Pennsylvania State College's war training program were given an experimental battery of tests. The

[1] Shortly after the entrance of the United States into the war, the program became officially known as Engineering, Science and Management War Training (ESMWT).

average age of this group was approximately eighteen and the average intelligence level as measured by the Otis S-A Test of Mental Ability was somewhat superior to that of college freshmen as found by Otis. The population samples used in the analysis of the several tests ranged in size from approximately 300 to 600 subjects. Each test was studied in relation to scores on the Cooperative Physics Test for College Students: Mechanics, Form 1936 B, which was administered to I.E.S. students at the close of their physics course. Three parts of the Revised Iowa Physics Aptitude Test were found to be closely related to attainment in physics as measured by the Cooperative Mechanics Test and were subsequently selected as sections of the Engineering and Physical Science Aptitude Test. These are the Mathematics, Formulation, and Physical Science Comprehension test parts.

Item analyses were made of the Moore Test of Arithmetic Reasoning, the Bennett Test of Mechanical Comprehension, Form AA, and the Otis S-A Test of Mental Ability, Higher Form A, the criterion again being scores on the Cooperative Mechanics Test. Eight items were selected from the Test of Arithmetic Reasoning on the basis of high biserial correlations with Cooperative Mechanics Test scores and a satisfactory range of difficulty. To this number were added two items believed to possess predictive value. The ten items, arranged in order of difficulty, comprise the Arithmetic Reasoning part of the Engineering and Physical Science Aptitude Test. Twenty-two items from the Bennett test were similarly selected on the grounds of item validity and range of difficulty and were developed into the Mechanical Comprehension test part.

An item analysis of the Otis test revealed that many highly verbal items bore substantial relations to the criterion. As a consequence, a selection was made of the most technical terms in the vocabulary section of the Moore-Nell Examination for Admission used at The Pennsylvania State College. These items comprise the test part entitled Verbal Comprehension.

The six parts which form the final compilation, their order of presentation and time limits are summarized below:

Number of Test Part	Title of Test Part	Time Limit
I	Mathematics	15 min.
II	Formulation	10 min.
III	Physical Science Comprehension	10 min.
IV	Arithmetic Reasoning	15 min.
V	Verbal Comprehension	10 min.
VI	Mechanical Comprehension	12 min.

The working time for the composite test is 72 minutes. Administration time, including test instructions, ranges from 80 to 90 minutes.

Interrelationship of Test Parts

Bingham has called attention to the fact that the probability of success in a college of engineering is determined by a variety of aptitudes. It has already been implied that a principal aim in the development of the Engineering and Physical Science Aptitude Test was to build a measure capable of appraising not one, but several kinds, of ability which underlie success in technical training. The obvious advantage of such an omnibus measure is that it opens the way to a more precise prediction of the intellectual complex under investigation by permitting the several test parts to be combined and optimumly weighted through a multiple regression technique. The maximum predictive efficiency of such a team of test parts ordinarily demands that each part be substantially correlated with the chosen criterion but that the several parts have low correlations with each other. The desirability of low intercorrelations among the several weighted parts of a test battery is clear. Inspection of Table 1 reveals that this expectation is moderately well founded for the Engineering and Physical Science Aptitude Test. The intercorrelations were computed from the test scores of 188 high school graduates receiving war training in Introductory Engineering Subjects.

Table 1

Intercorrelations of the Six Parts of the Engineering and Physical Science Aptitude Test (N = 188)

Test Part	Formulation	Phys. Sci. Comp.	Arith. Reasoning	Verbal Comp.	Mechanical Comp.	Median Intercorrelation
Mathematics	.609	.402	.544	.461	.290	.461
Formulation		.421	.445	.528	.328	.445
Physical Science Comp.			.475	.460	.588	.460
Arithmetic Reasoning				.405	.499	.475
Verbal Comprehension					.280	.460
Mechanical Comprehension						.328

The closest correspondence is represented by the correlation of .609 between the Mathematics and Formulation parts both of which are highly mathematical in content. The next highest figure, .588, is found for the relation between Physical Science Comprehension and Mechanical Comprehension. These two sections possess somewhat similar content, the former presenting the problems verbally and the latter pictorially. The Mechanical Comprehension section possesses the lowest median

intercorrelation with the other parts. With this exception, the median intercorrelations for the several test parts are very similar, ranging from .445 to .475. The median of all test part intercorrelations is .460.

Reliability

Coefficients of reliability were calculated for the composite test and for each of the six test parts by the split-half method. These data were based on the test scores of 201 men and women enrolled in engineering and science war training courses. The reliability coefficient for the composite test, corrected by the Spearman-Brown prophecy formula, was found to be .96. Reliability coefficients for the individual test parts are as follows: Mathematics, .90; Formulation, .68; Physical Science Comprehension, .87; Arithmetic Reasoning, .75; Verbal Comprehension, .93; and Mechanical Comprehension, .82.

Validation of Test

The Engineering and Physical Science Aptitude Test is designed primarily as a measure of training aptitude. As such, its validity must rest first upon its effectiveness in segregating those persons who are likely to progress rapidly in technical training from those who are likely to falter during such training. A second and more critical demonstration of the test's validity concerns its discriminatory power in industrial situations. Although both aspects of validity have been studied, this report is given over almost entirely to institutional validation, that is, to the predictive efficiency of the test in an academic situation. Brief mention will be made of findings obtained with the test in three industries where the criterion employed was again success during apprenticeship or training.

The test was administered to 188 Pennsylvania high school graduates taking the 1942 Summer Course in Introductory Engineering Subjects offered by The Pennsylvania State College. A large preponderance of these subjects were males. These students are given an intensive nine-week course in five technical subjects. Included in the program are engineering mathematics, engineering chemistry, engineering physics, engineering drafting and manufacturing processes. Final grades earned in each of the five subjects, an average final grade based on all subjects, and examination scores in three subjects were used as indices of training success. Statistical relationships between the total score and several part scores of the Engineering and Physical Science Aptitude Test and the various criteria were then established.

Table 2 presents the correlations obtained between composite and part scores on the aptitude test and final grades. The fact that many of

Table 2

Correlations between Scores on the Engineering and Physical Science Aptitude Test and Final Grades in Introductory Engineering Subjects (N = 188)

	Math. Grade	Chem. Grade	Physics Grade	Mfg. Proc. Grade	Draft. Grade	Average Grade
Eng. & Phys. Sci. Apt. Test (total score)	.590	.659	.717	.376	.346	.733
Part 1: Mathematics	.707	.585	.551	.137	.239	.626
Part 2: Formulation	.533	.595	.544	.200	.254	.587
Part 3: Phys. Sci. Comp.	.335	.482	.601	.380	.297	.561
Part 4: Arith. Reason.	.492	.321	.488	.260	.350	.522
Part 5: Verbal Comp.	.443	.570	.510	.320	.132	.542
Part 6: Mech. Comp.	.279	.322	.516	.359	.361	.485

these relationships are extremely high attests to the substantial validity of this test as a predictive instrument for this type of training. It is seen that total test score yielded a correlation of .73 with the average grade for the entire course. Correlations between the total score on the test and individual subject grades ranged from .72 for Physics to .38

Table 3

Correlations between Scores on the Engineering and Physical Science Aptitude Test and Examination Grades in Introductory Engineering Subjects (N = 188)

	Final Chem. Examination	Final Phys. Examination	Mid-term Draft. Examination (Visualization)	Mid-term Draft. Examination (Orthographic Projection)
Eng. & Phys. Sci. Apt. Test (total score)	.670	.712	.273	.324
Part 1: Mathematics	.571	.533	.057	.198
Part 2: Formulation	.528	.516	.190	.228
Part 3: Phys. Sci. Comp.	.501	.616	.292	.270
Part 4: Arith. Reason.	.362	.498	.318	.336
Part 5: Verbal Comp.	.609	.508	.131	.108
Part 6: Mech. Comp.	.325	.507	.382	.415

for Manufacturing Processes and .35 for Drafting. The field of drafting entails certain visualization and motor skills which are admittedly not sampled by the Engineering and Physical Science Aptitude Test. Hence, it is to be expected that the test is less closely related to this subject than to most of the others although selected parts of the test have since been

shown to afford a valuable index when supplemented by measures of the special skills involved in drafting work.[2]

Inspection of the table reveals that most of the test parts possess greatest predictive efficiency for those subjects with which they would seem logically to be most strongly associated. The Mathematics test part, for example, yields the highest correlation with mathematics grade; the Physical Science Comprehension test part correlates most highly with physics grade.

Table 4

Multiple Correlations between the Six Parts of the Engineering and Physical Science Aptitude Test Taken as a Battery and Grades in Introductory Engineering Subjects (N = 188)

Multiple Correlation between Six Parts of Aptitude Test and Final Grade in:	R	Coefficient of Forecasting Efficiency $1-\sqrt{1-r^2}$	Probable Error of Estimate $.6745\sigma\sqrt{1-r^2}$	Test Part Possessing Highest Weight
1	2	3	4	5
Mathematics	.732	.32	7.86	Mathematics
Chemistry	.731	.32	5.82	Mathematics
Physics	.738	.33	7.43	Physical Science Comprehension
Manufacturing Processes	.456	.11	6.95	Verbal Comprehension
Drafting	.429	.10	6.51	Mechanical Comprehension
Average Course Grade in All Subjects	.763	.35	4.54	Mathematics
Average Grade: Mathematics, Chemistry and Physics	.790	.39	5.63	Mathematics

It is seen in Table 3 that the aptitude test scores are related to final examination grades in much the same way that they are related to final grades in the several subjects. Again, the correlations are high with chemistry and physics criteria and low with drafting measures. It was not feasible to compute the relationships between the aptitude test and examination grades in mathematics and manufacturing processes since no uniform final examinations were given in these subjects.

When the six parts of the aptitude test are weighted by a multiple regression technique, the correlations with attainment in the several courses are raised. Reference to the multiple correlations in Table 4

[2] Griffin, C. H., Borow, H., and Reppert, H. C. Unpublished study on file. Personnel Service Division, The Pennsylvania State College, 1943.

will demonstrate this finding. The weighted team of test parts is seen to correlate most highly with an average composed of grades in mathematics, chemistry and physics, .790. The next highest multiple correlation, .763, resulted for Average Course Grade in all five subjects. The smallest multiple correlations were found for Manufacturing Processes, .456, and Drafting, .429. Column 3 presents the coefficient of forecasting efficiency when the weighted battery is used to predict grades in each subject. The probable error of a percentage-grade prediction for any subject is given in Column 4. For example, the average chance is one in two that a prediction of the average grade in the entire course will fall within 4.54 percentage points of the grade that is actually attained. Thus, chances are even that a student for whom an average course grade of 81.60 is predicted will earn a grade between 77.06 and 86.14. The test part receiving the highest partial regression coefficient in the multiple correlation with each grade criterion is given in Column 5.

Comparison of the foregoing figures with those reported for other tests falling under the broad classification of mechanical ability may serve to throw additional light on the relative merit of the Engineering and Physical Science Aptitude Test as a measure of aptitude for technical training. Bennett and Cruikshank have reported the results of numerous studies on the relationship between measures of mechanical ability and indices of success in engineering school. They list twenty group and individual tests for which are reported fifty-five correlations ranging from approximately zero to .91. The median correlation is .32 while only four correlations have a magnitude greater than .60. The only correlation which surpasses the .763 and .790 found in the present study is .91, listed for the Mann Dynamicube Test by Paterson, Schneidler and Williamson. This relationship obtained with sectioning scores in a drawing class. Other validity coefficients listed for the Dynamicube Test by both Bennett and Cruikshank, and Paterson, Schneidler and Williamson are .632 and .455 against grades in descriptive geometry, .56 against first-semester grades in mechanical drawing and .59 against first-year achievement in engineering school. Although differences in size, variability and nature of the population samples, as well as in the criteria utilized, do not warrant the drawing of rigorous mathematical inferences from a comparison of the foregoing correlational studies, it would seem reasonable to conclude that the Engineering and Physical Science Aptitude Test compares favorably with other existing instruments as a valid measure of aptitude for technical training on the college level.

Validity of Abbreviated Forms of Test

Inasmuch as the individual test parts of the Engineering and Physical Science Aptitude Test are not identical measures of the same func-

tion, they tend to contribute with varying degrees of importance to the prediction of the criterion. The practical implications of this condition are twofold: first, the test parts should be optimumly weighted by a regression method to effect maximum relationship with the predicted variable; this procedure, of course, was pursued in computing the multiple correlations presented in Table 4. Secondly, in the prognostication of most criteria, the test parts which make relatively minor contributions may be eliminated from the battery without appreciably adverse effect upon the accuracy of the prediction. The obvious advantage accruing from the use of an abbreviated form of the test is economy of time in test administration and scoring.

An empirical study has been carried out to investigate the practicability of short forms of the Engineering and Physical Science Aptitude Test.

Table 5

Comparison of Complete Form and Abbreviated Form of the Engineering and Physical Science Aptitude Test in the Prediction of Engineering Subject Grades (N = 188)

Criterion	Test Battery	R	Time Limit in Minutes	Test Battery	R	Time Limit in Minutes
Mathematics Grade	Complete Test	.732	72	Parts I, II, IV, V	.731	50
Mathematics Grade	Complete Test	.732	72	Parts I, IV	.718	30
Physics Grade	Complete Test	.738	72	Parts I, II, III, V, VI	.738	57
Physics Grade	Complete Test	.738	72	Parts I, III, VI	.712	37
Chemistry Grade	Complete Test	.731	72	Parts I, II, III, V	.717	45

Final grades in the introductory engineering courses were again used as criteria. An inspection was made of the partial regression coefficients deriving from the computation of the multiple correlations reported in Table 4. Those test parts possessing highest coefficients for the prediction of achievement in engineering mathematics, chemistry and physics were selected, and multiple correlations were computed between the selected test parts and each of the three forementioned criteria. The results are presented in Table 5. Examination of the table clearly shows that the abbreviated batteries possessed approximately the same degree of relationship with subject grades as the complete test, although the testing time required for the administration of each of the shorter batteries was considerably less than that required for the complete test. For instance, the first row reveals that a four-part weighted battery consisting of Mathematics (Part I), Formulation (Part II), Arithmetic Reasoning (Part IV) and Verbal Comprehension (Part V) yielded a multiple correlation of .731 with grades in engineering mathematics, this

figure being virtually identical with that obtained with the complete test. The time limit for the abbreviated battery, however, was 50 minutes, a saving of 22 minutes over the time limit of the complete test. When the battery was further reduced to include only the Mathematics and Arithmetic Reasoning (second row of Table 5), the multiple correlation dropped slightly to .718 whereas the time limit was pared to 30 minutes. It is evident, at least for the type of criteria employed in this study, that abbreviated forms of the Engineering and Physical Science Aptitude Test can frequently be employed with a substantial saving of time but with little effect upon the accuracy of prediction.

Studies with Women

The nation-wide recruitment of men for military service has heightened the problem of furnishing war industries with technically trained workers. The inevitable result has been that increasingly large numbers of women are being sought and trained for essential wartime occupations. Since many women do not adapt themselves to technical and mechanical tasks with the readiness of men, the training of unselected groups of women for such responsibilities is often slow and expensive. Numerous plants, as a consequence, have become seriously concerned with the problem of choosing those women among their applicants who reveal most promise of succeeding in technical apprenticeships.

Whereas the number of general ability tests suitable for women is great, few are available which have been standardized and validated for women in the engineering and related occupations. Recently, several studies have been conducted by the Personnel Service Division at The Pennsylvania State College to ascertain the value of the Engineering and Physical Science Aptitude Test as an index of women's aptitude for work entailing an understanding of technical principles and procedures. Only the briefest review of the findings may be presented here.

Fifty-three female high school graduates enrolled in introductory engineering chemistry courses were given the test. Total test scores were later correlated with grades in the mid-term course examination and a correlation of .65 was obtained. When the Mathematics, Formulation, and Verbal Comprehension test parts were selected and weighted as a three-factor team, the multiple correlation became .71.

The test was administered to 83 young women prior to laboratory training in quantitative chemical analysis at a large ordnance plant engaged in the manufacture of TNT. A criterion of performance accuracy was established by combining the ranks earned by each subject in titration accuracy, calculating accuracy and balance accuracy. The multiple correlation between the six parts of the Engineering and Physical

Science Aptitude Test and the composite rank in performance accuracy was found to be .41.

The Arithmetic Reasoning and Mechanical Comprehension test parts were used in a battery to predict the achievement of female drafting trainees at a large aircraft corporation. The battery also included the Minnesota Paper Form Board and the Thurstone Test of Surface Development. When scores in the four tests were equated for variability and combined into a single index of aptitude, a correlation of .85 between the composite index and achievement in training was found.

Three parts of the Engineering and Physical Science Aptitude Test were employed, among other measures, in a study designed to develop a valid battery for the selection of female chemical operator trainees at a synthetic rubber company. These parts were Physical Science Comprehension, Verbal Comprehension and Mechanical Comprehension. One criterion employed here was a work knowledge test comprising questions and problems on the principles and equipment emphasized during training. The test was administered upon completion of the training period. Correlations of the three parts of the Engineering and Physical Science Aptitude Test with the criterion were as follows: Physical Science Comprehension, .25; Verbal Comprehension, .49; Mechanical Comprehension, .34. The latter two parts were selected for inclusion in a five-variable prediction battery which yielded a multiple correlation of .65 with the criterion.

Two broad generalizations are suggested by the studies conducted with female populations:

(1) The Engineering and Physical Science Aptitude Test is significantly more difficult for women than for men. Concomitantly, several parts of the test would appear too exacting for women with a limited mathematics and science background or for women in training for positions which do not require a high level of technical understanding.

(2) In situations where women are receiving specialized training on an advanced level for work of an engineering and physical science nature, the test is of appreciable value as a prediction measure.

Standardization

Percentile norms for the Engineering and Physical Science Aptitude Test were established on the basis of test results of men and women entering engineering and science courses. These persons were part of a larger group enrolled in the Engineering, Science and Management War Training Program of The Pennsylvania State College. A descriptive study of a sample of this larger population indicated that the age of the average enrollee was approximately 30. Eighty-five per cent of

the subjects had completed high school training and a substantial proportion of these had continued their education beyond the high school level. The median salary was between thirty and forty dollars per week. Approximately half of the enrollees held professional, skilled, and clerical positions and only about ten per cent of those employed were unskilled workers. It is apparent that the group was somewhat superior educationally, economically, and, in all probability, mentally, to a general, unselected population. This description is thought to be reasonably accurate for the standardization groups which comprised a large proportion of the ESMWT population studied.

Norms for men are based on 6695 cases; norms for women are based on 2295 cases. In both instances, percentile norms were established for composite test scores. A table of test means and standard deviations for high school graduates as well as information on the functions of the test in one ESMWT program may also be found in the test manual (4).

Summary

The Engineering and Physical Science Aptitude Test was developed to assess suitability for training in technical work on the high school graduate or college level. It has been employed extensively as an aptitude measure for persons enrolled in engineering and science courses in the ESMWT program of The Pennsylvania State College. Each of the six parts which comprise the test measures a relatively independent aspect of technical performance. Percentile norms have been established separately for 6695 men and 2295 women enrolled in war training courses.

Studies on the validity of the test were carried out on high school graduates enrolled in the Introductory Engineering Subjects curriculum of The Pennsylvania State College's war training program. Grades in engineering mathematics, chemistry, physics, manufacturing processes and drafting served as criteria. Multiple correlations between the weighted battery of test parts and course achievement were as high as .79. Reference to previous studies on the engineering school validity of numerous other tests indicated that the prediction efficiency of the Engineering and Physical Science Aptitude Test is among the highest reported in the testing literature.

In the prediction of success in technical training, the test parts which contribute least to the forecast can, in many cases, be eliminated without perceptibly decreasing the strength of the relationship between the test and the criterion of achievement during training. When multiple correlations were obtained with abbreviated forms of the test whose parts were carefully selected, little effect upon the accuracy of prediction resulted although the saving of test administration time was considerable.

The Engineering and Physical Science Aptitude Test is distinctly more difficult for women than for men. Nevertheless, in situations where women are in training for positions demanding a high level of technical ability, the test has been demonstrated to possess high discriminating value. Studies carried out both with female engineering war trainees receiving classroom instruction and with several types of female industrial apprentices have revealed close agreement between test performance and training success.

Received August 29, 1943.

References

1. Bennett, G. K. *Test of mechanical comprehension, form AA.* The Psychological Corporation, New York, 1940.
2. Bennett, G. K., and Cruikshank, R. M. *A summary of manual and mechanical ability tests.* The Psychological Corporation, New York, 1942.
3. Bingham, W. V. *Aptitudes and aptitude testing.* Harper and Brothers, New York and London, 1937.
4. Borow, H. *Manual for engineering and physical science aptitude test.* The Psychological Corporation, New York, 1943.
5. Farwell, H. W., Lapp, C. J., et al. *Cooperative physics test for college students: Mechanics; form 1936B.* Cooperative Test Service, New York.
6. Griffin, C. H., Borow, H., and Reppert, H. C. *Predicting the achievement of women in engineering drafting training.* Unpublished study on file. Personnel Service Division, The Pennsylvania State College, 1943.
7. Moore, B. V. Analysis of results of tests administered to men in engineering defense training courses. *J. appl. Psychol.*, 1941, **25**, 619–635.
8. Moore, B. V. *Test of arithmetic reasoning.* The Psychological Corporation, New York, 1941.
9. Moore, B. V., Lapp, C. J., and Griffin, C. H. *Engineering and physical science aptitude test.* The Psychological Corporation, New York, 1943.
10. Moore, B. V., and Nell, R. B. *Examination for admission.* The Pennsylvania State College, State College, Pennsylvania.
11. Paterson, D. G., Schneidler, G. G., and Williamson, E. G. *Student guidance techniques.* New York: McGraw-Hill Book Co., Inc., 1938.
12. Stuit, D. B., and Lapp, C. J. Some factors in physics achievement at the college level. *J. exper. Educ.*, 1941, **9**, 251–253.
13. Stuit, D. B., and Lapp, C. J. *Revised physics aptitude test; Iowa placement examinations, new series, revised.* Bureau of Educational Research and Service, State University of Iowa, Iowa City, Iowa.

A Study of the Reliability of the Minnesota Rate of Manipulation Test by the Split-half and Test-retest Methods

Jacob Tuckman

Jewish Vocational Service, Cleveland, Ohio

The Minnesota Rate of Manipulation Test[1] is used to select workers for office and factory jobs where speed of hand and finger manipulation is important. The apparatus is a wood board containing 60 cylindrical holes, arranged in four rows of fifteen into which 60 slightly smaller blocks can be placed. The test consists of two parts: Placing—in which the subject, using one hand, places the blocks into the holes in a definite order from a fixed position; and Turning—in which the subject picks up a block with the left hand, turns it over, and puts it back into the same hole with the right hand, alternating hands for each subsequent row. One practice and four test trials are given. The score is the number of seconds required to complete the four test trials.

The purpose of this study is to determine the reliability of the Minnesota Rate of Manipulation Test since the manual accompanying the test does not present any reliability data. The split-half and test-retest techniques were used.

Test scores were available for 386 men, 319 women, 145 boys and 111 girls, who were referred for testing by the Counseling and Placement Departments of the Jewish Vocational Service in Cleveland. The men ranged in age from 17 to 56 years with a mean age of 21.9; the women from 17 to 45 years with a mean age of 21.8. Approximately one-half of the adult group were high school graduates. The mean percentile rank for intelligence as measured by the Pressey Senior Classification Test and the Otis Higher Examination was 73 for men and 52 for women. The boys ranged in age from 13.4 to 18.9 years with a mean age of 15.8; the girls from 13.9 to 18.4 with a mean age of 15.6. All were attending school. The grade range was from 9B to 12A. The median grade was 10B (first half of the tenth year) for both boys and girls. The mean percentile rank for intelligence as measured by the American Council On Education Psychological Examination for High School Students and College Freshmen (1938, 1939, and 1940 editions) and the Terman Group Test of Mental Ability was 78 for boys and 75 for girls.

[1] Developed by W. A. Zeigler and distributed by The Educational Test Bureau, Inc., Minneapolis, Minnesota.

The reliability coefficients for Placing and Turning, together with the means (time in seconds), and the standard deviations for all four groups are presented in Table 1. The coefficients were obtained by correlating trials 1 and 3 and trials 2 and 4 and then applying the Spearman-Brown prophecy formula. The corrected coefficients which range from .91 to .97 indicate that the reliability of both Placing and Turning is high. There is no difference between the adult and school group nor is there any sex difference.

Table 1

Odd-even (Trials 1 and 3 and Trials 2 and 4) and Corrected Reliability Coefficients, Means (Time in Seconds), and Standard Deviations for Placing and Turning for Men, Women, Boys, and Girls. N = 961

		Placing				Turning			
	N	r*	Corr. r†	Mean[1]	S.D.	r*	Corr. r†	Mean[1]	S.D.
Men	386	.87	.93	224.0	21.2	.90	.95	172.0	20.8
Women	319	.93	.96	227.0	23.3	.84	.91	176.9	24.2
Boys	145	.90	.95	236.3	19.9	.94	.97	183.4	24.3
Girls	111	.89	.94	236.8	22.1	.91	.95	183.6	22.5

* P.E. r equals .01 for all groups.

† Corrected by Spearman-Brown prophecy formula.

[1] For discussion regarding differences in performance between adult and school groups, and sex differences in the adult group, see J. Tuckman, "A comparison of norms for The Minnesota Rate of Manipulation Test," *J. appl. Psychol.*, 1944, **28**, 121–128.

An additional 100 high school students—51 boys and 49 girls—were tested to determine test-retest reliability. The group ranged in age from 13.3 to 18.5 years with a mean age of 16.3 years. The median school grade was 10A (the last half of the tenth grade). Intelligence ratings as measured by the American Council on Education Psychological Examination for High School Students and College Freshmen (1940 and 1941 editions) were available. The mean percentile rank for intelligence was 81. The interval between the initial test and the final test varied from 1 to 14 days, with a median interval of 7 days.

The split-half and corrected reliability coefficients of the various trials on the initial and final tests for both Placing and Turning are presented in Table 2. In comparing one trial with another, or trials 1 and 3 with trials 2 and 4, there is little difference between the initial and final tests. The implications of these findings are significant especially for the industrial situation where time limitations present a real problem. Norms developed on the basis of the number of seconds taken to complete the first two trials would have satisfactory reliability. For Placing, the corrected reliability coefficients between trials 1 and 3 and trials 2 and

4 are .92 and .93 for initial and final tests, respectively; for Turning, it is .94 for both. However, the test-retest reliability coefficients drop to .82 for Placing and .81 for Turning. The correlation between score on Placing and score on Turning on the initial test is .60; on the retest it is .57 which is identical with the correlation reported by Zeigler for a group of 500 subjects.

Table 2

Split-half and Corrected Reliability Coefficients of the Various Trials on the Initial and Final Tests for Placing and Turning for High School Students. N = 100

	Placing						Turning					
	Initial Test			Final Test			Initial Test			Final Test		
Trial	r	P.E. r	Corr. r^*	r	P.E. r	Corr. r^*	r	P.E. r	Corr. r^*	r	P.E. r	Corr. r^*
1 and 2	.87	.02	.93	.78	.03	.88	.81	.02	.90	.79	.03	.88
1 and 3	.80	.02	.89	.80	.02	.89	.82	.02	.90	.84	.02	.91
1 and 4	.70	.03	.82	.77	.03	.87	.76	.03	.86	.73	.03	.84
2 and 3	.92	.01	.96	.85	.02	.92	.84	.02	.91	.82	.02	.90
2 and 4	.85	.02	.92	.79	.03	.88	.80	.02	.89	.78	.03	.88
3 and 4	.85	.02	.92	.83	.02	.91	.85	.02	.92	.84	.02	.91
1+3 & 2+4	.86	.02	.92	.87	.02	.93	.89	.01	.94	.88	.02	.94

* Corrected by Spearman-Brown prophecy formula.

The mean score (time in seconds) and the standard deviation for trials 1, 2, 3, 4 and for all 4 trials combined on the initial and final tests for Placing and Turning are presented in Table 3. There is a progressive though small decrease in time required for each successive trial for both

Table 3

Mean Score (Time in Seconds) and Standard Deviation for Trials 1, 2, 3, 4 and for All Four Trials Combined on the Initial and Final Tests for Placing and Turning for High School Students. N = 100

	Placing				Turning			
	Initial Test		Final Test		Initial Test		Final Test	
Trial	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
1	59.4	6.1	54.2	4.6	48.0	5.9	40.1	4.7
2	58.1	4.9	53.5	4.9	45.8	5.3	39.3	4.3
3	57.3	5.3	53.4	4.6	45.1	5.2	39.0	3.9
4	56.8	4.7	53.2	4.8	44.8	5.3	38.6	4.0
1, 2, 3, 4 combined	232.0	20.0	214.2	17.2	183.8	20.3	157.4	15.3

Placing and Turning on the initial and final test. The largest decrease occurs between trial 1 on the final test and trial 4 on the initial test, and is greater for Turning than for Placing.

The gains in total score made on the final test over the initial test for both Placing and Turning are noteworthy. Without exception the performance of all subjects was faster on the final test for both Placing and Turning. These differences are statistically significant. The $\frac{D}{\sigma \text{ diff.}}$ between total score on initial and final test is 6.7 for Placing and 10.4 for Turning. In comparing the performance of the Cleveland Group with the norms of The Educational Test Bureau, the mean score for Placing on the initial test is equivalent to the 52nd percentile; the mean score on the final test is equivalent to the 79th percentile. For Turning the mean score on the initial test is equivalent to the 63rd percentile;[2] the mean score on the final test is equivalent to the 92nd percentile. On the basis of our observations of the test performance of these subjects, the practice effect appears to be the result of less tension and greater familiarity with test procedure rather than an increase in native speed of hand or finger manipulation.

Norms based on the performance of these 100 high school students for initial and final tests for Placing and Turning are presented in Table 4.

Table 4

Norms for Initial and Final Tests for Placing and Turning. 100 High School Students

	Placing		Turning	
Percentile	Initial Test	Final Test	Initial Test	Final Test
99	177.5	173.8	145.5	125.5
90	210.0	193.8	158.2	139.5
80	214.7	199.9	164.8	144.5
70	220.7	204.3	172.2	149.0
60	227.3	209.2	179.1	153.1
50	231.7	214.3	183.1	156.5
40	236.0	218.4	187.8	159.8
30	240.2	222.8	193.5	164.0
20	246.6	227.5	199.2	168.9
10	259.5	234.5	212.7	177.3
1	281.5	263.5	233.5	202.5

Summary

1. The internal consistency of The Minnesota Rate of Manipulation Test is high. Although the test-retest reliability coefficients are not

[2] For discussion regarding the tendency of subjects to perform more rapidly on Turning than on Placing see J. Tuckman, "A comparison of norms for The Minnesota Rate of Manipulation Test," *J. appl. Psychol.* 1944, **28**, 121–128.

quite as high as those obtained by the split-half method, the reliability of both Placing and Turning appears to be satisfactory for individual diagnosis.

2. In the test manual the author indicates that when a person has done poorly in one trial, that trial may be discarded and an additional one given since the substitution of a good trial would not materially affect the score. This may hold for the substitution of a single trial. However, if a subject is retested at least within a 2 week interval the norms of The Educational Test Bureau are definitely not applicable.

Received August 13, 1943.

A Preliminary Report on Some New Tests of Musical Ability*

Robert William Lundin

Indiana University

The purpose of the present investigation was to devise music tests which would measure directly and in an objective fashion those aspects of musicianship which are commonly taught in musical theory courses. By this we mean such abilities as those resulting from ear training, harmonization, and dictation.

The present tests do not purport to measure any innate capacity or abilities. This does not mean, however, that before the onset of formal training everyone would do equally well on the present set of tests or any other music tests. Since musical ability is the result of a complex of many factors interacting in the individual's life history, informal training may begin early in life. The child may be brought up in a musical environment having many musical experiences, therefore being constantly subjected to music. In grade school he attends music classes where he may learn the bare fundamentals of singing and the appreciation of music. Interest and biological equipment may also be important contributing factors. Thus, by the time students have arrived at the college level, some are more equipped than others to succeed in a musical career. Consequently, some will do better on music tests.

The Tests

The present battery consists of five tests which were recorded in the Indiana University radio studios. In each test subjects were given two practice trials before starting the test proper. The number of each item was announced on the record before the subject heard the music which he was to judge. In all cases S was to make a response of "same" or "different." The items were arranged in random order so that a subject would get no more correct than he deserved after the chance factor had been corrected for. There were an equal number of "s" and "d" items.

The first test which we have called interval discrimination, consisted of twenty-five pairs of step intervals, each interval consisting of two notes. For each item there were two sets of step intervals. *S* was asked to tell whether or not the second *interval* was the same as or different from the first. When an interval was the same, the number of steps or

* Grateful appreciation is hereby acknowledged to Dr. J. B. Carroll for his careful direction of this investigation.

notes on the scale which lay between the first and second note of each interval was the same. However, this does not mean that the notes would be the same since they were not, but in some cases the intervals were.

The second test, melodic transposition, consisted of twenty-five pairs of single melodies. The second melody was *always* in a different key from the first. Sometimes the transposed melody was not the same as the first, there being changes of one note so that if the second melody were transposed back into the key of the first, it would be different. If *S* considered the second melody really the same as the first (except for the change in key), he responded by an "s," but if he detected a change in the melody, he marked a "d."

Test three, harmonic transposition, consisted of thirty-four sets of single chords and resolutions with the second item of each set appearing in a different key from the first. The transposed item was either the same as the first in *harmonic structure* or different from it. The test began with a series of single chords, then sets of two chords and finally three. In the case of single chords, a chord was played in one key, then another. In order for the subjects to indicate an "s" response, the harmonic construction of the two chords had to be alike. For example, two major chords in the same position would elicit an "s" response. However, if the first chord were a major chord and the second chord a minor triad, then the correct response would be a "d." When two and three chords were given the test became similar to test two. A maximum of three chords was used because it became extremely difficult to remember more than three chords for purposes of comparison.

The fourth test was called melodic sequences. Groves' *Dictionary of music and musicians* defines a sequence as ". . . the repetition of a definite group of notes or chords in different positions of the scale like regular steps ascending or descending."[1] This test consisted of twenty sequential groups. Each item contained four such groups or sequences. In all cases the first three patterns were correct since they followed a definite order, but sometimes the last group did not follow the same pattern as the first three. In such cases the subject would respond with a "d." If the sequence seemed correct he would reply with an "s." All sequences were diatonic and began and ended in the key of C.

Test five, harmonic sequences, was similar to test four. It consisted of eighteen sets of *modulating harmonic* sequences. Here *S* was asked to judge whether the last group of the sequential pattern followed the same pattern as the preceding ones.

The tests were given to two groups of subjects. In the first group, there were 23 music students from the DePauw University School of

[1] Groves, *Dictionary of music and musicians.* New York: Macmillan Co., 1938.

Music and 37 students from the Indiana University School of Music. The second group consisted of 100 students taken from various undergraduate classes in psychology at Indiana. The latter group we shall call the unselected group since they were taken without consideration as to whether or not they had had any training or interest in music. The group selected from the two music schools will be known as the musicians since most of them were preparing for careers as professional musicians or music teachers.

Summary of the Statistical Analysis

1. For purposes of reliability both groups were used making N = 160. Although the reliabilities of the separate tests were not high the total scores are fairly reliable (.73). If unreliable items and items that are too difficult were eliminated, the revised tests should result in higher reliabilities for the separate tests.

Table 1

Correlations between Test Scores and Sum of Ratings

Test	*r*
Interval Discrimination	.46
Melodic Transposition	.51
Harmonic Transposition	.26
Melodic Sequences	.37
Harmonic Sequences	.29
Total Test Scores	.53

Table 2

Correlations of Various Test Scores and Ratings of Teachers

Test	Melodic Dictation	Harmonic Dictation	Written Harmonization	General Ability
Interval Discrimination	.55	.45	.39	.47
Melodic Transposition	.61	.51	.43	.28
Harmonic Transposition	.36	.13	.42	.36
Melodic Sequences	.58	.52	.30	.36
Harmonic Sequences	.44	.38	.26	.30
Total Test Scores	.47	.48	.52	.59

2. The tests were validated on the musicians alone since no adequate criteria were available for the unselected group. The teachers were asked to rate their students on a graphic rating scale devised by the author for this purpose. These ratings of professors were used as criteria in validating the test. Students were rated on the following traits: (a) melodic dictation, (b) harmonic dictation, (c) written harmonization, and (d) general ability in the theory course. The general ability rating

was based on the students' grades and daily exercises. The correlations between ratings and test scores appear in Tables 1 and 2. Tests of interval discrimination and melodic transposition correlated highest with the sum of the ratings. Interval discrimination, melodic transposition, melodic and harmonic sequences correlated highest with instructors' ratings of students in melodic and harmonic dictation. The test of harmonic transposition correlated the lowest with total ratings.

None of the correlations are extremely high but validity coefficients seldom are. We can see that the tests *do* appear to correlate with the teachers' ratings on abilities taught in theory courses.

3. The object of giving the test to a group of unselected subjects was to see if there was a significant difference between the performance of the musicians and the unselected subjects on the various tests and to see if the test discriminates between trained and untrained subjects. When the means of the musicians and the unselected group were compared, a very significant difference was found on every test as well as on the total scores. When the means of the DePauw and Indiana musicians were compared no significant differences were found.

When the means of the unselected students who indicated an indifference toward classical music were compared with those who liked classical music, a difference was found which was to some extent statistically significant. However, when the means of those who liked classical music and those who liked it very much were compared, a significant difference was found. Therefore, we would expect also a very significant difference between those who were indifferent toward classical music and those who liked it very much.

4. Indiana and DePauw professors tended to rate their students about the same.

5. By item analysis, we found that there are certain items in each test which discriminate between musicians and the total group. The items in each test need to be arranged according to difficulty and those items which when corrected for chance give a .00 proportion passing should be eliminated.

6. To what extent the present battery of tests will predict success remains to be seen. Right now they are in the early stages and will have to undergo careful revision before any final conclusion can be drawn. One thing seems clear, the tests appear to measure ability to discriminate intervals, recognize melodies and harmonies when they are transposed, and ability to detect mistakes in harmonic and melodic sequential patterns. This type of musical behavior, the author believes, is characteristic of a musical personality and, therefore, will tend to discriminate a musician from one who is not trained in this type of activity.

Received August 18, 1943.

Leadership as Related to the Bernreuter Personality Measures: IV. An Item Analysis of Responses of Adult Leaders and Non-Leaders

Nelson G. Hanawalt and Helen M. Richardson*

New Jersey College for Women, Rutgers University

In the preceding paper (9) of this series adult male leaders and non-leaders are compared as to their average scores on the Bernreuter and Flanagan personality scales. Significant group differences are found on some of the scales. The purpose of the present study is to pick out the individual items on the Inventory on which the leaders and the non-leaders show significantly different responses. We are also interested in determining the relative weight of the various "traits" of the Inventory which appear to be responsible for the difference in response. When the leaders and the non-leaders differ in their response to an item, what traits are responsible for this difference? To a lesser extent we are interested in the validity of the Scale traits and the difference in response of the different kinds of leaders.

In a previous paper (5) the authors report a similar investigation comparing college leaders and non-leaders. In this study, based upon 36 leaders and 45 non-leaders, it is found that there are twelve items which show a significant difference in response between the two groups. There are twelve additional items which have P values between .05 and .10. An analysis of four of the six Scales of the Inventory (Self-Sufficiency, Introversion, Dominance, and Sociability) by the average weight difference method, described below, reveals that dominance is the most heavily contributing factor to the obtained difference on the twelve significant items. Extroversion and self-sufficiency contribute equally, whereas there is little difference in sociability. This is what we would expect from the comparison of the gross averages on the Inventory as a whole, excepting in the increased difference in the case of self-sufficiency (8). That there is no difference between the two groups in sociability appears to be due to the fact that Flanagan's trait of sociability is not simple and unitary gregariousness, though statistically it may be an independent factor. It is also found that the leaders make significantly fewer "?" responses.

* The authors are indebted to Miss Mary J. McNaughton and Miss Betty Steiner for assistance in computing the Chi-Squares in one section of the results, and to the National Youth Administration for clerical assistance.

Subjects and Method

Leadership, as described in the preceding article (9), is investigated in its vocational and social aspects, i.e., as found in Supervisors and Office-Holders. These two kinds of leadership are there fully defined. Also there is a full description of our sample of subjects, method of selection, etc., consequently these data will not be repeated here. In the present paper the responses of 90 Supervisors are compared with those of 88 Non-Supervisors, and 57 Office-Holders with 70 Non-Office-Holders. These subjects are the same as used in the previous paper excepting that here only the first 70 of the original 116 Non-Office-Holders are used in order to cut down on the amount of tabulation.

The responses of the four groups of subjects on the 125 items of the Inventory were tabulated. The items which showed considerable difference in the distribution of responses between the leaders and the non-leaders were evaluated for significance by the application of the Chi-Square test (4, p. 385 f.). In this way a list of items were identified on which the Supervisors differ significantly from the Non-Supervisors, and a second list on which the Office-Holders differ significantly from the Non-Office-Holders.

Table 1

Distribution of Responses and Chi-Square Values for Items with P Values of .05 or Less for 90 Supervisors and 88 Non-Supervisors

Items (P = .01 or less)		Responses Yes	No	?	x^2
*111. Have you been the recognized leader (Pres., etc.) of a group in the last 5 years?	(S)	65	23	2	17.986
	(nS)	37	50	1	
* 27. Do your interests change rapidly?		6	79	5	14.722
		25	60	3	
* 91. Does your mind often wander so badly that you lose track of what you are doing?		8	81	1	14.548
		28	59	1	
78. When you are in low spirits do you try to find someone to cheer you up?		24	57	9	12.700
		42	45	1	
* 69. Do you often find that you cannot make up your mind until the time for action has passed?		14	75	1	11.983
		31	53	4	
* 54. Do you often feel lonesome when you are with other people?		14	75	1	11.983
		31	53	4	
* 88. If you came late to a meeting would you rather stand than take a front seat?		28	56	6	11.962
		50	34	4	
* 36. Have you ever solicited funds for a cause in which you were interested?		70	20	0	11.195
		48	39	1	
98. Do you prefer a play to a dance?		71	14	5	9.999
		50	29	9	
74. Do you ever upbraid a workman who fails to have your work done on time?		56	32	2	9.673
		35	47	6	
(P lies between .02 and .01)					

Table 1.—*Continued*

43. Do you like to bear responsibilities alone?	74	14	2	8.906
	55	27	6	
85. Do you consider the observance of social customs and manners an essential aspect of life?	75	12	3	8.892
	56	26	6	
117. Do you usually take added responsibilities on yourself?	64	25	1	8.824
	49	30	9	
(P lies between .05 and .02)				
123. Does discipline make you discontented?	18	66	6	7.506
	34	49	5	
80. Do you lack self-confidence?	12	77	1	7.078
	18	63	7	
* 49. Do you often experience periods of loneliness?	18	71	1	6.739
	31	54	3	
* 71. Do you experience many pleasant or unpleasant moods?	35	46	9	6.611
	51	32	5	
79. Can you usually understand a problem better by studying it out alone than by discussing it with others?	60	24	6	6.555
	42	36	10	
58. Do you ever complain to the waiter when you are served inferior or poorly prepared food?	61	29	0	6.337
	45	41	2	
95. Would you "have it out" with a person who spread untrue rumors about you?	65	22	3	6.242
	75	9	4	
18. Are you touchy on various subjects?	31	57	2	6.189
	43	40	5	
87. Do you take the responsibility for introducing people at a party?	54	32	4	6.179
	39	38	11	
48. Does it bother you to have people watch you at work even when you do it well?	16	73	1	6.036
	29	57	2	

Note: The *number of responses* of the Supervisors under each category appear in the first row after each item and the number of Non-Supervisors' responses in the second row. The starred items are significant for both the Supervisor and the Office-holder comparisons and are referred to in the text as overlapping items.

In addition to finding the particular items which show a significant difference, the relative contribution of the "traits" to the obtained difference were also evaluated. This method consists in finding for each item the average weight difference of the leaders and non-leaders on the various Scales of the Inventory. Thus if there is a large average weight difference between the leaders and the non-leaders on an item of the Inventory in dominance, for instance, and little difference in sociability, we can conclude that dominance is a factor contributing to the found difference whereas sociability is not. The validity of this analysis is limited by the validity of the Scale weights.

Item Analysis of Responses of Leaders and Non-Leaders

The application of the Chi-Square Test to the two groups of leaders vs. the non-leaders reveals that there are 23 items in each case which show

a significant difference (P value of .05 or less). However, only 9 of the 23 items are common to both lists. These lists of significant items are found in Tables 1 and 2, together with the number of *S*'s making each kind of response and the Chi-Square values. The items which are common to the two lists are starred. The distribution of responses for the Office-Holders (Table 2) are in per cent in order that the two groups of *S*'s, which are of different sizes, may be more easily compared with respect to their responses. Of course the Chi-Squares are based upon the actual number of responses in each category.

Table 2

Distribution of Responses and Chi-Square Values for Items with P Values of .05 or Less for 57 Office-Holders and 70 Non-Office-Holders

Items (P = .01 or less)		Responses Yes %	No %	? %	χ^2
* 36. Have you ever solicited funds for a cause in which you were interested?	(O)	88	12	0	28.814
	(nO)	41	56	3	
35. Would you dislike any work which might take you into isolation for a few years, such as forest ranging, etc.		60	32	9	13.170
		41	59	0	
40. Have you ever organized any clubs, teams, or other groups on your own initiative?		77	21	2	13.088
		49	51	0	
*111. Have you been the recognized leader (pres., etc.) of a group within the last five years?		95	5	0	12.543
		70	30	0	
* 69. Do you often find that you cannot make up your mind until the time for action has passed?		12	88	0	12.081
		31	61	7	
56. Are you careful not to say things to hurt people's feelings?		86	12	2	10.518
		60	33	7	
96. At a reception or tea do you feel reluctant to meet the most important person present?		7	91	2	10.051
		21	69	10	
8. Do you often feel just miserable?		23	75	2	9.865
		44	49	7	
46. Do jeers humiliate you even when you know that you are right?		12	84	4	9.819
		34	59	7	
121. Do you like to be with people a great deal?		75	18	7	9.243
		50	41	9	
* 49. Do you often experience periods of loneliness?		18	79	4	9.232
		43	54	3	
(P lies between .02 and .01)					
* 27. Do your interests change rapidly?		7	89	4	8.835
		27	69	4	
* 71. Do you experience many pleasant or unpleasant moods?		44	53	4	8.674
		60	29	11	
* 91. Does your mind often wander so badly that you lose track of what you are doing?		7	91	2	8.529
		27	71	1	
103. Do you have difficulty in starting a conversation with a stranger?		11	88	2	7.989
		31	67	1	

Table 2.—*Continued*

125. Would you feel very self-conscious if you had to volunteer an idea to start a discussion among a group of people? (P lies between .05 and .02)	7 24	91 73	2 3	7.825
20. Do you feel self-conscious in the presence of superiors in the academic or business world?	19 34	81 60	0 6	7.761
11. Do you try to get your own way even if you have to fight for it?	51 29	39 61	11 10	7.260
45. Have you ever had spells of dizziness?	23 34	77 59	0 7	7.120
* 54. Do you ever feel lonesome when you are with other people?	11 29	88 69	2 3	6.652
73. Does some particularly useless thought keep coming into your mind to bother you?	14 26	81 60	5 14	6.535
* 88. If you came late to a meeting would you rather stand than take a front seat?	37 59	60 40	4 1	6.099
94. Do you ever take the lead to enliven a dull party?	75 54	19 36	5 10	6.086

Note: This table is to be read the same as Table 1, excepting that the *distribution of responses is in per cent.* This is done in order to facilitate comparisons of the responses of Office-holders and Non-Office-holders.

In the case of the Supervisors vs. the Non-Supervisors there are 13 items where the difference is very significant (P is .02 or less). In the case of the Office-Holders vs. the Non-Office-Holders there are 16 such items. This is further evidence, as is indicated in the preceding paper, that the Office-Holders differ more markedly from the Non-Office-Holders than the Supervisors do from the Non-Supervisors.

On the overlapping items in the two lists (Tables 1 and 2) it is found that the distribution of responses for each item is about the same for the two groups of leaders vs. non-leaders, i.e., the proportion of "Yes," "No," and "?" responses are quite similar in the two tables. It is interesting, however, to inquire about the distribution of responses on the non-overlapping items. How do the Office-Holders respond to an item which does not differentiate significantly between them and their contrasting group but does differentiate the Supervisors from the Non-Supervisors, and vice versa? In each table there are 14 of these non-overlapping items. We shall take the case of the Supervisors first. Of the 14 non-overlapping items on their list, the following items show the same trend of response for the Office-Holder comparison but the difference falls a little short of significance: 98, 85, 117, 123, 80, 58, 95, 18, and 87. On the remaining five items, however, there is a marked difference in the distribution of responses on the Office-Holder comparison. It is on these items where we see most clearly the difference between the

Supervisors and the Office-Holders. More Office-Holders seek someone to cheer them up when in low spirits, are less likely to upbraid a workman for not having work done on time, do not like as well to bear responsibilities alone, can solve problems better by discussing them with others, and are a little less likely to be bothered by having people watch them at work. These differences indicate that the Office-Holders are less self-sufficient and more gregarious than the supervisors.

In the case of the items on the Office-Holder comparison (Table 2) which do not overlap with the Supervisors, we find that the Supervisors show a similar trend in response on the following items: 35, 56, 8, 46, 125, 45, 73, and 94. On the other six items there is a marked difference in the distribution of responses. The Supervisors have organized fewer clubs, etc., feel more reluctant to meet the important people at tea, have less desire to be with people, have more difficulty in starting a conversation with a stranger, are more likely to feel self-conscious in the presence of superiors, and are less likely to fight for their own way. Here again the chief difference between the Supervisors as a group and the Office-Holders as a group is that the Office-Holders are less self-sufficient and more sociable. There is also an indication in both the above comparisons that the Office-Holders are also more dominant. These are also differences which are indicated in the preceding paper (9).

In our study of college leaders it was found that the leaders made significantly less use of the "?" response than the non-leaders. This was interpreted to mean that leaders make decisions more easily than non-leaders. In the present study we again find that there is a tendency for the leaders to use the "?" response less frequently, but the difference is not so clear cut as in the case of the college study. On the list of significant items (Table 1) the Office-Holders encircle the "?" in 3% of their responses compared with 5.5% for the Non-Office-Holders. This is not a reliable difference. The Supervisors use this response in 3.4% of their responses compared with 5.4 for the Non-Supervisors. This is likewise an unreliable difference. On the Inventory as a whole we find the same trend.

An Analysis of the Contribution of Scale Traits

In addition to knowing that the leaders and the non-leaders respond differently to an item, it is interesting and valuable to know something of the personality factors which are responsible for the difference in response. The average weight difference method reported in our study on college leaders (5) furnishes the means of making such an analysis within the limits of the Scales of the Inventory. In the college study the analysis was based upon four Scales: Self-Sufficiency, Introversion, Dominance,

and Sociability. In the present analysis Neurotic Tendency is added. It is true that some of these scales have a high correlation with each other, but this is often not apparent on an individual item.

The method of average weight difference consists in finding the difference in average weight gained by the leaders and the non-leaders on each item and for each Scale used in the analysis. The number of "Yes," "No" and "?" responses for each item is multiplied by the weight given for that response. These products are added algebraically and divided by the number of *S*'s. The average weight difference between the leaders and the non-leaders is secured by subtraction. These differences on the various Scales can be compared as to their relative size. In Table 3 we

Table 3

Average Weight Difference of Office-Holders vs. Non-Office-Holders on the Neurotic Tendency Scale

(Item No. 8)

Office-Holders: (N = 57)

Response	n		Weight		
Yes	13	×	7	=	91
No	43	×	−7	=	−301
?	1	×	0	=	0
				Total	−210

−210 ÷ 57 = −3.684 (Average weight)

Non-Office-Holders: (N = 70)

Yes	31	×	7		217
No	34	×	−7	=	−238
?	5	×	0	=	0
				Total	− 21

−21 ÷ 70 = −.300 (Average weight)

Average Weight Difference = −3.384 in favor of Office-holders

have reproduced one of these operations to set forth more clearly this process. A computation like this example is made for each of the five Scales on each item listed in Tables 1 and 2. Rather than print a long and uninteresting table containing these data, the results are set forth graphically in Figures 1 and 2. The average weight difference of the example in Table 3 is graphed in Figure 2 under item No. 8. It will be noted that this difference in Neurotic Tendency is by far the largest difference on this item. The difference between the Office-Holders and the Non-Office-Holders on this item, "Do you often feel just miserable?", is largely due to the fact that the Non-Office-Holders are more neurotic. The Office-Holders are more extroverted and slightly more dominant and self-sufficient on this item. There is no difference in sociability.

It is interesting to note the difference in the relative size of neurotic tendency and extroversion on the above item. Despite the high correlation between these two Scales, the difference in response on this item is much more a matter of neurotic tendency than it is introversion. Inspection of Figures 1 and 2 will show other items where this difference is also

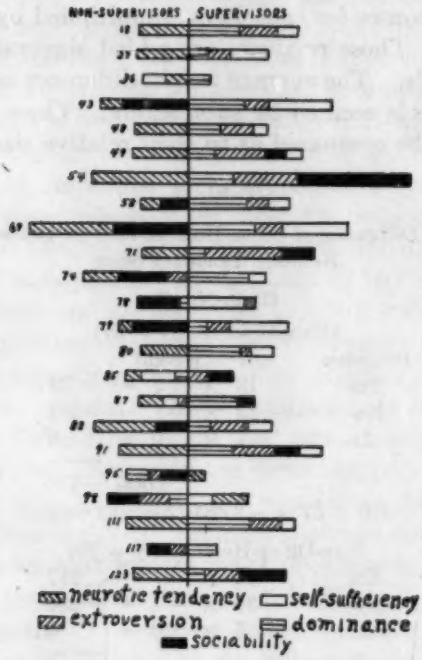

FIG. 1. Average weight difference for supervisors and non-supervisors. If the average difference is in favor of the leaders on a scale it is plotted on the right side of the vertical line; if it is in favor of the non-leaders it is plotted on the left side. The length of the space alloted to a particular trait on a bar is proportional to the size of the average difference between the leaders and the non-leaders on that trait.

quite marked. The Introversion-Extroversion Scale is plotted in terms of extroversion, hence this scale and Neurotic Tendency are usually found on opposite sides of the graph. In Figure 1, on items 74 and 87 they are both on the same side. This means that on these two items the Non-Supervisors are both more neurotic and more extroverted, a reversal from what we would expect from the high correlation between the Scales. Of course the graphs give only the relative difference of the various Scales. If there are any who are interested in the absolute size of the difference, it can be closely approximated by taking the length of the sample bar in the legend as one-half a point. In the case of a bar in

which a particular Scale is not represented, it can be concluded that there is negligable difference on that Scale.

Space does not permit an analysis of each item in Figures 1 and 2. The items are reproduced in Tables 1 and 2. By referring to the corresponding bar in the graphs each item can be analyzed as we have done

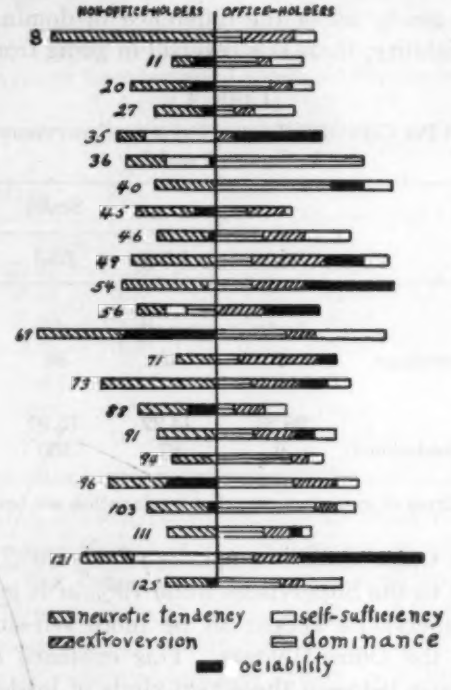

FIG. 2. Average weight difference for office-holders and non-office-holders. This figure is to be read the same as Figure 1.

for item No. 8 above. It is apparent from a general perusal of the graphs that the Office-Holders differ more from the Non-Office-Holders than the Supervisors do from the Non-Supervisors. The total amount of difference is about one-third greater in the case of the Office-Holders. From the relative amount of space taken in the two graphs by the various Scales, it is apparent that Neurotic Tendency and Dominance are the ones which account for the greatest amount of Total Difference. The exact amounts of these Total Differences are found in Table 4.

A comparison of the Total Difference, however, does not tell us anything concerning the distribution of these differences. They are pictured in the graphs as falling on the leader or the non-leader side. The per cent of the Total Difference on each Scale favoring the leaders is also recorded

in Table 4. Thus in the case of B1-N only 7% favors the Supervisors, i.e., 93% of the difference in neurotic tendency is in favor of the Non-Supervisors, while all of the difference on this Scale is in favor of the Non-Office-Holders. For Self-Sufficiency 86% of the difference favors the Supervisors and 67% the Office-Holders. B3-I is plotted in terms of extroversion and most of the difference is in favor of the leaders. The leaders also have nearly all of the difference in dominance. However, in the case of Sociability, there is a reversal in going from the Supervisor

Table 4

Total Difference and Per Cent of Difference Favoring Supervisors and Office-Holders in Figures 1 and 2

	Scales				
	B1-N	B2-S	B3-I	B4-D	F2-S
Supervisors					
Total Diff.	19.16	12.12	13.29	16.56	14.34
Per cent favoring Supervisors	7	86	84	91	39
Office-holders					
Total Diff.	28.23	14.22	15.90	24.06	16.02
Per cent favoring Office-holders	0	67	100	98	79

Note: B3-I is in terms of extroversion. For explanation see text.

comparison to the Office-Holder comparison: only 39% of the difference in sociability goes to the Supervisors while 79% of it goes to the Office-Holders. The Supervisors appear to be more self-sufficient and less extroverted than the Office-Holders. This evidence corroborates the personality difference between these two kinds of leadership which was pointed out above in connection with the non-overlapping items.

In view of the fact that the Supervisors appear to be more clearly differentiated from the Non-Supervisors in Figure 1 than was the case in the previous comparison (9) based upon the Inventory as a whole, we decided to make a test of the significance of the difference in average scores on the basis of the significant items listed above. In order to cut down on the labor of such a task we use only 34 subjects in each group. These are in no way selected but just happen to be the number whose responses were recorded on one tally sheet. These data are found in Table 5. Two comparisons are included from the Office-Holder group which did not show a reliable difference in the former study. The critical ratios of the former paper are also included in this table for convenience. It will be noted that all of the critical ratios for this abbreviated list of items are larger excepting in the case of Sociability. The greatest change

is in the case of self-sufficiency for the Supervisor comparison. This difference is reliable here whereas in the former comparison it was not. The same is true in the case of dominance although the critical ratio in the former study shows a high probability of a real difference. It is strange that the reliability of the differences in sociability should both decrease in the light of the general increase on the other Scales.

Table 5

Difference in Average Score between Leaders and Non-leaders Based upon the 23 Significant Items. (N = 34)

Scales	Mean	S.D.	Mean	S.D.	C.R.	C.R. on Inventory as a Whole
	(Supv.)		(Non-Supv.)			
B1-N	−22.97	18.25	−1.03	20.55	−4.70	−2.94
B2-S	13.41	9.27	4.68	11.13	3.52	1.09
B4-D	21.59	16.80	4.82	15.15	4.30	2.72
F2-S	6.38	10.30	5.44	14.80	0.30	1.37
	(O-H.)		(Non-O-H.)			
B2-S	8.88	8.35	3.79	8.55	2.48	−0.18
F2-S	−6.50	9.75	−1.47	10.60	−2.03	−2.33

Since the data are available in Table 5 we could not resist the temptation to test the significance of the differences in average scores of the Supervisors vs. the Office-Holders on self-sufficiency and sociability. The difference in average in the former case is 4.5 and the C.R. is 2.12. Had all of the cases been used this might have turned out to be a reliable difference. On Sociability, however, there is a significant difference. The average difference is 12.88 and the C.R. is 4.90, the Office-Holders being more sociable.

Discussion

There are some facts which stand out throughout this whole series of studies: the leaders are better adjusted (less neurotic, more extroverted, and more self-confident) and they are more dominant. This fact comes out in the comparisons of the leaders vs. the non-leaders on the Inventory as a whole, and with even more emphasis in the analysis based upon the significant items. Examination of the significant items listed above reveals that they are essentially questions of adjustment and dominance with only a few referring to sociability or self-sufficiency. We have gone over the items on the two lists in the light of Bernreuter's (1) and Flanagan's (3) most significant items, and other evidence, and have rated each one as characteristic of one of the above traits. On the nine overlapping

items, six are essentially neurotic tendency, and three dominance. Considering the Supervisor list as a whole, there are nine each for neurotic tendency and dominance, and five for self-sufficiency or sociability. Considering the Office-Holder list as a whole there are eleven which are essentially neurotic tendency, ten dominance, and two sociability or self-sufficiency. A completely objective method of bringing out the same point is to count the number of our significant items which overlap with Bernreuter's (1) and Flanagan's (3, p. 46 f.) more diagnostic items. Combining the items of our two lists we have a total of 37 different items. Seven of these appear on one or more of the three adjustment scales, seven on the dominance list, and only two on the self-sufficiency or sociability lists.

In this series of investigations there has been consistently an unreliable difference in self-sufficiency in comparing the average scores of our contrasting groups on the Inventory as a whole. Yet in the more detailed analyses it appears quite certain that the Supervisors are more self-sufficient than the Non-Supervisors. There is strong indication that this is also the case in the Office-Holder comparison and in the college leadership comparison (5). The Inventory as a whole masks this difference between the leaders and the non-leaders, apparently because Scale self-sufficiency is confused on a number of items with withdrawal. There is a healthy self-sufficiency which can be differentiated from withdrawal. This distinction is not well focused in the Scale. The leaders are more self-sufficient but on the Inventory as a whole this is masked by the self-sufficiency credit which the non-leaders receive for the tendency to withdraw because of an inferior adjustment. St. Clair and Seegers criticize this Scale from a somewhat similar point of view as a result of their empirical studies with college students (6, 7).

In an item analysis of college leaders vs. non-leaders (5), we criticize Scale sociability as not being identical with gregariousness. In the above study it proved enlightening to examine each of the significant items in the light of the weights given to the different answers on the sociability scale. It was found that the non-leaders were often rated more sociable on an item because of their more frequent use of the "?" response, or because they appeared to be more sensitive to the presence and opinions of other people. The "?" response often has a very heavy sociability weight. It is difficult to see the connection here with gregariousness. We have presented evidence to show that this response is more closely tied up with an inferior adjustment. Also the usual picture of gregariousness is one of self-confidence in the presence of others.

In the present study when both the Supervisors and the Office-Holders are rated more sociable on an item it is because they are better

adjusted. For instance, both groups of leaders are rated more sociable than the non-leaders because they are not as likely to feel lonesome with others, they do not have periods of loneliness, or many pleasant and unpleasant moods, and are not as apt to have a wandering mind. On the items where the Office-Holders alone are rated as more sociable it appears to be due to genuine sociability. However the telling evidence against the sociability scale is brought out by an examination of the items on which the *non-leaders* are rated as more sociable. Both the Non-Supervisors and the Non-Office-Holders are rated as more sociable because they have difficulty in making up their minds or are not sure about it, and they would rather stand than take a front seat if they came late to a meeting. In addition the Non-Supervisors are rated as more sociable because they do not like to bear responsibilities alone or are not sure about it, are not likely to upbraid a workman or are not sure about it, they try to find someone to cheer them up when in low spirits, can understand a problem better by discussing it with others, prefer a dance to a play, and do not like to take on added responsibilities or question it. The Non-Office-Holders are rated as more sociable because they feel reluctant to meet the important people at tea, are not sure about having dizzy spells, and are not sure about feeling self-conscious in the presence of superiors. When the non-leaders are rated as more sociable on an item it is often because of the heavy sociability weight given to the "?" response. In the case of the Non-Supervisors it is often due to the sociability weight given to the lack of self-sufficiency and dominance. With the exception of a few items which show sociability in its usual meaning, the confusion which arises in this Scale in the present study seems quite definitely to be due to the fact that sometimes a healthy adjustment receives a sociability rating and sometimes the opposite is true.

The idea of St. Clair and Seegers of using the Bernreuter Scales as a profile appears to have possibilities in the case of leadership. The typical profile of a leader would appear to be low on Neurotic Tendency, high on Dominance, and above average on Self-Sufficiency. Since the inclusion of the non-significant items often blurs the picture, perhaps a better scheme for picking potential leaders would be to select the leadership discriminating items and weight them in accordance with the reliability of the difference between the contrasting groups. It seems quite certain that by such a system we should be selecting people with good adjustment, dominant, and self-sufficient, but would not have to trouble ourselves about Scale traits. Naturally such a test would be open to all the objections of the pencil and paper test of this type, but it would be superior to the Inventory as a whole in selecting potential leaders.

Summary

An item analysis by means of the Chi-Square Test was made to determine the items on the Inventory which show a significant difference in responses of the leaders vs. the non-leaders: Supervisors are compared with Non-Supervisors and Office-Holders with Non-Office Holders. There are twenty-three items in each comparison which show a significant difference in response, nine of which are common to the two lists.

It is found that the leaders differ most from the non-leaders on the Adjustment Scales (Neurotic Tendency, Introversion-Extroversion, and Self-Confidence) and in Dominance. There is no reliable difference on the Sociability scale. Although there is no reliable difference on the Scale as a whole in Self-Sufficiency, there is a reliable difference between the Supervisors and the Non-Supervisors on the significant items in this respect. The reliability of the difference between the leaders and the non-leaders on all of the Scales is increased by comparing them on only the significant items excepting in the case of Sociability.

Comparing the two groups of leaders (Supervisors and Office-Holders) on the significant items, it is found that the Office-Holders are more Sociable than the Supervisors, and that the Supervisors have a high probability of being more Self-Sufficient than the Office-Holders. The Office-Holders also appear to be more Dominant. On the whole the Office-Holders appear to differ more sharply from their contrasting group than do the Supervisors.

It is concluded on the basis of the results of this paper, and on the results of this series of studies as a whole, that the chief difference between the leaders and the non-leaders is one of adjustment: the leaders are better adjusted in general than the non-leaders. The leaders are also more dominant but this difference is not altogether independent of adjustment. It also appears that the leaders are more self-sufficient (especially the supervisors) but this difference does not show up on the inventory as a whole because of the fact that the non-leaders receive Self-Sufficiency credit for a tendency to withdraw which masks the difference between the two groups in this respect. That there is no difference in Sociability between the leaders and the non-leaders is apparently due to the fact that sometimes a healthy adjustment receives Sociability credit and sometimes the opposite is true.

It is suggested that a leadership test, free from the confusion of Scale traits, could be constructed on the basis of the significant items of the Inventory for predicting leadership.

Received August 29, 1943.

References

1. Bernreuter, R. G. The theory and construction of the Personality Inventory. *J. soc. Psychol.*, 1933, **4**, 387–405.
2. Bernreuter, R. G. *Manual for the personality inventory.* Stanford: Stanford University Press. Pp. 6.
3. Flanagan, J. C. *Factor analysis in the study of personality.* Stanford: Stanford University Press, 1937. Pp. 103.
4. Garrett, H. E. *Statistics in psychology and education.* New York: Longmans, Green, 1937. Pp. 493.
5. Hanawalt, N. G., Richardson, H. M., and Hamilton, R. J. Leadership as related to the Bernreuter personality measures: II. An item analysis of responses of leaders and non-leaders. *J. soc. Psychol.*, 1943, **17**, 251–267.
6. St. Clair, W. F., and Seegers, J. C. Certain aspects of the validity of the Bernreuter Personality Inventory. *J. educ. Psychol.*, 1937, **28**, 530–540.
7. St. Clair, W. F., and Seegers, J. C. Certain aspects of the validity of the F-Scores of the Bernreuter Personality Inventory. *J. educ. Psychol.*, 1938, **9**, 299–315.
8. Richardson, H. M., and Hanawalt, N. G. Leadership as related to the Bernreuter personality measures: I. College leadership in extracurricular activities. *J. soc. Psychol.*, 1943, **17**, 237–249.
9. Richardson, H. M., and Hanawalt, N. G. Leadership as related to the Bernreuter personality measures: III. Leadership among adult men in vocational and social activities. *J. appl. Psychol.*, 1944, **28**, 308–317.

Simplifying the Scoring Technique of the Bernreuter Personality Inventory

Homer Kempfer

New York State Education Department

If the Bernreuter Personality Inventory is used extensively, the scoring process involves either considerable labor or a cash outlay and a time delay. This investigation was made in an effort to find a simpler scoring technique resulting in a minimum loss of reliability.

The Personality Inventory was administered in the regular manner to one hundred twenty boys in grades IX to XII inclusive. Stagner's[1] short-cut method of scoring was used to obtain the regular scores. These raw scores were converted into percentile ranks by using the tentative percentile norms furnished by the Bernreuter manual.[2]

Because each response on any item of the blank may have a value anywhere on a fifteen-point scale, any attempt to reduce this variability would be expected to make the discrimination of the instrument less accurate. Nevertheless five such simplifications were tried. Strip cardboard keys for each of the four scales were made as follows:

Key I. All original values of 3 to −3 inclusive were assigned a value of 0; 4 and above became 1; −4 and below became −1. Scoring time per blank for the four scales was thereby cut to approximately two minutes.

Key II. Original values of 2 to −2 inclusive were assigned a value of 0; 3 and above became 1; −3 and below became −1. This key on four traits required approximately three minutes to apply.

Key III. Question mark responses were ignored and all "Yes" and "No" values having a difference of seven or more were paired. Positive values became 1 and negative values became −1. On scale B3-I this required difference was reduced to five in order to retain sufficient items to yield a discriminating final score. Scoring time was less than two minutes per blank.

Key IV. Question mark responses were ignored and all positive items were counted as 1 if there was a difference of six or more between the "Yes" and "No" answers. This difference was reduced to five on scale B3-I for the same reason as in Key III. Scoring time was eighty seconds per blank.

Key V. All positive responses were counted as 1. Zero and negatives were ignored. Scoring time was less than two minutes per blank.

[1] Stagner, R., Scoring weighted item tests: A short-cut method. *Person. J.*, 1932, 11, 174–176.

[2] Bernreuter, R. G., *Manual for the Personality Inventory.* Stanford, Calif.: Stanford University Press, 1935, p. 4.

After the new scores were derived, Pearson correlations were worked for each scale with results from each key. The coefficients are shown in Table 1.

Table 1

Correlations of the Five Trial Keys with Original Percentile Ranks on the Four Scales. N = 120 High School Boys

Scale	Key I	Key II	Key III	Key IV	Key V
B1-N	.870	.693	.877	.884	.853
B2-S	.820	.871	.838	.821	.785
B3-I	*	.832	.811	.834	.861
B4-D	.852	.906	.823	.870	.900

* Omitted because of extremely small number of items remaining to yield a score.

The coefficients with four exceptions all fall in the "eighties" with .84 as a central tendency. Probable errors all are within reasonable limits, ranging from $\pm.011$ to $\pm.032$.

Conclusions and Recommendations

1. No new key tried was sufficiently meritorious to replace the original Bernreuter keys for accurate work.

2. For quick location of extreme cases the best new key for each trait could be used. Individuals in the middle fifty percent as determined by the new keys could be considered "normal." Cases in either extreme twenty-five percent could be rescored with the regular Bernreuter keys. For quick grouping of cases the loss in reliability is believed justified by the saving of more than 90% in scoring time.

3. Further investigation should be made using a larger population both to determine more definitely the better keys and to furnish adequate percentile norms.

4. Similar investigation should be made with the Flanagan Scales.

Received August 25, 1943.

Simplified Scoring of the Bernreuter Personality Inventory*

David C. McClelland

Wesleyan University

The Bernreuter Personality Inventory (1, 2) is one of the most popular questionnaires for measuring personality traits. Super's recent survey of it (7) has underlined the extremely wide use it has had, both as a research and as a personnel tool. In the latter capacity it has often served at colleges as a means of classifying large groups of entering Freshmen with respect to the six traits measured. It has been used for this purpose at Wesleyan University, and has proven valuable in picking out the extreme cases of maladjustment that will probably be in need of counselling.

One of the chief limitations of the test, however, is that it takes so long to score. If the hand keys are used, one person with an adding machine needs from six to ten minutes to score the test for one trait. If a group is to be tested, consisting of, for example, around 200 men, the size of an entering class at Wesleyan, 100 to 150 hours of tedious labor are required to score the tests for all six traits. The test may be scored more quickly by the IBM machine, a procedure recently developed by Helmick (5) working at Wesleyan, but even then each answer sheet must be run through the machine three times for each trait, and the scores obtained must be combined. The process is still very time-consuming compared with the ease with which most achievement tests are scored. Also, there remains the difficulty that for individual testing machine scoring is not always available, or if it is, so much time is needed to change the scoring keys to run the answer sheet through three times that little is saved over the tedious hand scoring.

In view of these facts, it would be extremely desirable to have a much simpler and quicker method of scoring the Inventory, if one could be devised that would correlate highly with the full score. One way of simplifying the scoring would be to do away with the weights assigned to each answer. Peterson and Dunlap (6) have recently thrown doubt on the need for the weighting process by showing that if all the weights in the Strong Vocational Interest Blank are reduced to 1, 0, and -1, scores may be obtained that correlate around .97 with the total score for most vocations. Another way of simplifying the scoring would be to

* I should like to record my gratitude to Miss Alice Donahue for her valuable clerical assistance in the preparation of this report.

reduce the number of answers which contribute to the total score for a given trait. Many of the answers on the Personality Inventory are not important for one trait (though they may be for another) and are weighted only ±1 or 0. Experience in scoring indicates that, as a rule, these answers of minor importance tend to cancel each other out and influence the net result very little. Consequently it was decided to simplify and shorten the scoring of the Inventory (a) by omitting entirely from the score for a given trait, any answer to a question which was not weighted at least ±3, and (b) by weighting all such answers simply +1 or −1. To put it another way: The number of answers weighted −3 or more by Bernreuter which were chosen by a subject were simply *counted* and subtracted from the number of answers weighted +3 or more chosen by the subject.

For example: Item No. 1 on the Inventory is weighted +2, −2, and 0 for a "Yes," "No," or "?" answer on the B1N scale. In the simplified scoring on this scale, no answer on this question would be counted. On the other hand, Item No. 2 is weighted +5, −4, and −2 for "Yes," "No," and "?" on B1N. In the simplified scoring a "Yes" answer would be scored +1, a "No" answer −1, and a "?" answer 0 or not at all. Table 1 shows the number of answers scored +1 and the number scored −1 for each of the five traits investigated.

Table 1

Number of Answers for Each Bernreuter Trait Scored Usually and in the Simplified Method. Total possible = 375

	Simplified Scoring			Usual Scoring
Trait	Weighted +1	Weighted −1	Total	Total
B1N	54	43	98	344
B2S	35	34	69	341
B4D	36	38	74	344
F1C	88	57	145	307
F2S	53	97	150	310

B3I was omitted because it correlates so highly with B1N (2). The figures make it clear that for any given trait only from one-quarter to one-half of the answers usually scored are included in the shortened scale. A larger proportion of the Flanagan scales, F1C and F2S, is included in the short form since the weights he assigned are as a rule heavier.

The shortened method of scoring was applied to the Inventories of a group of 114 students whose full Bernreuter scores for all traits had already been calculated.[1] The resulting values were correlated with the

[1] I am greatly indebted to Mr. John S. Helmick who made these scores available for my use.

full scale values. Table 2 presents the correlations obtained for each trait. They are all .95 or above except for the B2S scale. On the chance that such high correlations might be due to some peculiarity of the group of subjects used, another correlation between the two methods of scoring the B1N scale was calculated for a second group of 108 college men. It was .963. Table 2 also presents the reliability coefficients for each of the scales as published by Bernreuter (2). The comparison shows that

Table 2

Correlations between Full and Simplified Scoring (r_{fs}, N = 114) Compared with Reliability Coefficients of Full Scores (r_{ff}, N = 70)

	B1N	B2S	B4D	F1C	F2S
r_{fs}	.966	.844	.955	.975	.955
r_{ff}	.91	.92	.89	.86	.78

the shortened score is at least as good an index of the total score as the total score is of itself on a second occasion. This means that no one should worry that a shortened score will be off by a few percentiles any more than he worries that any full score will be off by a few percentiles from what it would be if the subject took the test over again immediately. In every trait except B2S the short score may be substituted for the full score with little loss in accuracy. The reason for the correlation being so much lower in the case of B2S is not known. On this trait, users of simplified scoring may either run the risk of being off by a number of percentiles in more cases, or else drop out the B2S scale altogether, substituting in its place the F2S scale which yields much the same information.

There remains the problem of interpreting a given short score in terms of some norms. This may be done in two ways. (1) The short score may be directly converted into a full score, and the published Bernreuter norms for college men used. Equations for this purpose for each trait are presented in Table 3. They are derived from the usual formula

Table 3

Equations for Converting a Given Short Score (s) to Its Most Probable Full Score Equivalent (f) in a Population of College Men

B1N	f = 5.6s	−23.7
B2S	f = 5.1s	+ 4.0
B4D	f = 5.5s	+ 1.3
F1C	f = 4.9s	−15.7
F2S	f = 3.8s	+ 9.1

(cf. 4, p. 73) for predicting one score from a correlated score. In working out the equations in Table 3, the correlations were rounded off to two places and the means and standard deviations to one place, following

Table 4

Tentative Percentile Norms for the Bernreuter Personality Inventory Simplified Scoring

Note: N = number of cases; M = mean of distribution of scores; SD = standard deviation of the distribution of scores.

Raw Score	College Men B1N N = 301	B2S N = 300	B4D N = 211	F1C N = 177	F2S N = 212
40			99		
35			99	99	99
30	99	99	97	99	99
25	99	99	93	97	99
20	97	96	84	95	98
18	96	95	82	93	96
16	95	90	76	92	96
14	92	86	69	90	95
12	90	82	62	86	92
10	89	73	58	84	90
8	84	66	50	82	86
6	82	58	42	79	84
4	76	46	35	76	82
2	73	38	31	69	76
0	66	31	24	66	73
− 2	62	21	18	62	69
− 4	54	16	16	58	62
− 6	50	12	12	54	58
− 8	46	7	8	50	54
−10	38	5	6	42	46
−12	35	3	4	38	42
−14	27	2	3	35	38
−16	24	1	2	31	30
−18	21		1	27	27
−20	16		1	21	24
−22	14			18	18
−25	10			16	14
−30	5			10	8
−35	2			6	5
−40	1			3	2
−45				1	1
M =	−6.0	4.5	8.1	−7.3	−9.2
SD =	14.2	8.6	11.5	16.9	14.8

Bernreuter's practice. The equations make it very simple to convert an obtained short score into a full score which may then be looked up in Bernreuter's published percentile norms. They should not be used, however, unless the subject tested comes from a population not demonstrably different from that of college men. The means and standard

deviations for college women, for example, would be quite different and the transformation equation would consequently be altered also.

The short scores need not be converted. They may be interpreted in terms of the tentative percentile norms presented in Table 4. These norms are based on the scores made at entrance by Wesleyan college men in 1938, 1939, and 1943. They include the 114 cases used in obtaining the correlation reported in Table 2. The percentile values are approximate since they are derived from the standard scores used at Wesleyan by means of a table of equivalent values (4, p. 60). They should, of course, be used only for college men at entrance, but practically they may be applied without too much inaccuracy to cases belonging to population groups, the means and standard deviations for which, as published by Bernreuter (3), do not differ greatly from those he reports for college men. It is interesting to note that the percentile equivalent of a given short score in Table 4 is very similar to what it is if the short score is converted to a full score and looked up in Bernreuter's norms, even though the standardizing groups are different.

The short scores correlate highly with the full Bernreuter scores. They are easily interpreted in either one of the two ways just described, but above all, they reduce enormously the amount of work necessary to score the Inventory. In machine scoring, keys may be made out so that any answer weighted +3 or more is counted as a Right, and any weighted −3 or more as a Wrong. The total score may then be obtained by running the answer sheet through the machine only once on the Right-Wrong formula. In hand scoring, the simplest method seems to be to take a thin, translucent sheet of paper, place it over the first page of the Inventory, and mark +, in ink, over those answers weighted +3 or more, and − over those answers weighted −3 or more. The sheet may be turned over and a key made similarly for the second page on its back, but a new sheet is necessary for the third page. These keys are then placed over a used Inventory; the number of − answers encircled are counted and subtracted from the number of + answers encircled. The time taken for scoring a single Inventory for one trait in this way is reduced from around 8 minutes to around 1 minute, and all the troublesome addition is avoided.

Summary

The Bernreuter Personality Inventory may be scored by counting all answers chosen that are weighted +3 or more and subtracting all answers chosen that are weighted −3 or more, for each trait scale. The resulting short scores correlate with the full scores above .95 for B1N, B4D, F1C, and F2S, and .84 for B2S. For interpretation, the short score may be converted directly to the full score by formulae given, or it may be placed

in the tentative short score percentile norms for college men, also supplied. The simplified scoring greatly reduces the time and labor necessary for either hand or machine scoring of the Personality Inventory, and consequently greatly increases the efficiency and usefulness of the test.

Received September 15, 1943.

References

1. Bernreuter, R. G. The theory and construction of the personality inventory. *J. soc. Psychol.*, 1933, **4**, 387–405.
2. Bernreuter, R. G. *Manual for the personality inventory.* Stanford University: Stanford University Press, 1938.
3. Bernreuter, R. G. *The personality inventory: tentative percentile norms.* Stanford University: Stanford University Press, 1938.
4. Greene, E. B. *Measurements of human behavior.* New York: Odyssey Press, 1941.
5. Helmick, J. S. Changes during college attendance in scores on Bernreuter's Personality Inventory. Unpublished Master's thesis, Wesleyan University, 1942.
6. Peterson, B. M., and Dunlap, J. W. A simplified method for scoring the Strong Vocational Interest Blank. *J. consult. Psychol.*, 1941, **5**, 269–274.
7. Super, D. E. The Bernreuter Personality Inventory: a review of research. *Psych. Bull.*, 1942, **39**, 94–125.

An Empirical Evaluation of Study Habits in Elementary German*

J. R. Wittenborn
Yale University
and
R. P. Larsen
University of Illinois

Students, teachers, and educational counselors are constantly challenged by the problem of selecting or recommending effective methods of study. In response to this challenge, numerous excellent study manuals for the student and pedagogical papers for the teacher and the counselor have been prepared. With a few important exceptions, the value of the study procedures suggested in these useful publications has never been systematically ascertained.

There is a limited number of excellent general suggestions for study. Aside from their direct application, they serve valuably in creating desirable attitudes and in stimulating the student to an aggressive systematic approach to his own problems. In many instances, however, general suggestions are of limited value to the bewildered undergraduate seeking assistance in study procedures. Much of the published material contains hints which are expressed in general terms only; few of the suggestions are specific to an area or course of study.

In counselling the student who is academically inefficient, an examination of the study methods that he employs is usually in order. Correct diagnosis of study difficulties and faults is not always easy. It is necessary to know what specific procedures are of the greatest probable value, and it is important to be able to identify those procedures or characteristics which are deleterious.

The empirically determined value of specific characteristics or of procedures in the study of a particular course or area would offer the counselor assistance in the diagnosis. Unfortunately, however, no direct approach to this problem appears to be feasible and the writers have been obliged to employ an indirect approach which involves the use of a questionnaire procedure.

* The writers wish to acknowledge the assistance of the Personnel Bureau and the German Department of the University of Illinois. Assistance in tabulating the data was granted by the W. P. A. under the supervision of the University of Illinois.

In the present study an attempt has been made to evaluate forty-two items in a study habit questionnaire for students in the first semester of college German. The study habit items included in the questionnaire were those describing particular procedures, attitudes, or difficulties. In compiling the list, an effort was made to include study characteristics which were remediable. Specific questions concerning remediable practices or situations were devised in order to provide maximum diagnostic value for the questionnaire and in order to permit direct remedial application.

Most of the items are not original in content. The ideas in many were borrowed from other writers; some came from teachers; and others were inspired by an earlier investigation.[1] They were expressed as questions requiring the mutually exclusive 'yes' or 'no' type of answer.

The Plan of the Investigation

Two criteria were employed in validating the items:

1. The total score earned on the Cooperative German Achievement test Elementary Form 0.
2. Course marks earned in the study of elementary German.

In December, 1940, the questionnaire shown in Table 1 was administered to all of the classes in the first semester of German. The Cooperative German test was given to all of the classes in January.

The marks earned in the elementary German course for the first semester, 1941, were taken from the files of the Liberal Arts and Science College. The students whose records were included for study were Freshmen in the College of Liberal Arts and Science at the University of Illinois during the academic year of 1940–41. None of them had had previous experience with a Germanic language.

In evaluating the items on the basis of scores earned on the Cooperative German test, an attempt was made to hold intelligence constant. The measure employed for intelligence was the total score earned on the American Council on Education Psychological Examination. This test was a part of the guidance battery which was administered to all Freshmen entering the College of Liberal Arts and Science in September, 1940.

The Treatment of the Data

The results of the two validating procedures, as well as a statement of the items, are given in Table 1.

[1] Larsen, R. P., Wittenborn, J. R., and Giesecke, E. G. Factors contributing to achievement in the study of first semester college German. *J. exper. Educ.*, 1942, 10, 265–271.

Table 1

This is not a test; it is an inventory of your study habits and attitudes. You are asked to respond to the following list of statements, not according to how you think you should study, but *according to how you actually do study.* Please answer all the statements. Your frankness in answering these items will be very important to us in our attempt to find those study habits which are desirable. Your answers will *not influence your grades* in any way. Please place a (+) in the parenthesis in front of each statement to which you would answer "yes" and a (−) in front of each to which you would answer "no." These statements refer only to the study of *German.* Answer every statement.

	Correlation of item with:			Number of students in each grade classification				
				A	B	C	D	E
				30	42	47	29	10
	German Achievement	Psychological Examination	German Achievement minus Psychological Examination	Per cent of students in each grade classification who marked item (+)				
				A	B	C	D	E
	1	2	3	4	5	6	7	8
1. I study German at a time of day when I work efficiently.	−.13	.07	−.22	.77	.81	.68	.72	.70
2. While I am studying, I worry about whether I will remember the material.	−.09	.00	.00	.23	.35	.40	.41	.70
3. I find it hard to keep my mind on what I am studying.	−.11	.14	−.26	.28	.43	.38	.41	.60
4. Talking, noises, etc., severely handicap me while I am studying.	.44	−.32	.87	.63	.48	.51	.48	.40
5. I often put off or evade the study of my language assignment.	−.13	−.12	−.07	.23	.38	.38	.41	.40
6. When writing examinations, I get so nervous that I can't write down all I know.	−.31	−.32	−.15	.27	.26	.38	.55	.40
7. I often get tense and nervous when I study.	−.12	.07	−.20	.13	.21	.15	.24	.40

8. While studying, I frequently get up, walk around, glance at a paper, or otherwise interrupt my work.	.11	.13	.05	.60	.36	.57	.48	.20
9. I am able to distinguish between what should and what should not be memorized by rote.	.08	−.31	.37	.77	.74	.72	.59	.50
10. I understand what the instructor wants to achieve by having the students do certain things in class.	.25	−.47	.77	.97	.91	.83	.76	.80
11. I practice and drill all new material in order to make it as nearly automatic and habitual as possible.	.13	−.45	.58	.57	.43	.36	.38	.40
12. I look upon each lesson as a step leading toward mastery of the language.	.09	−.15	.24	.87	.69	.53	.65	.80
13. I want to learn as much German as possible during the time that I will be taking it.	.26	.28	.12	1.00	.95	.81	.83	1.00
14. I take each assignment as it comes, regarding it as just another tedious task.	.08	−.04	.13	.20	.36	.30	.45	.40
15. I try to master each lesson or assignment.	.44	.35	.30	.90	.86	.64	.62	.80
16. I follow a study schedule which allows me ample time for German preparation.	−.14	−.30	−.16	.33	.40	.36	.38	.50
17. I prepare my assignments by going through them only once.	.24	.41	.03	.33	.43	.21	.14	.20
18. I nearly always study my German assignment before the next meeting of the class.	−.04	−.11	−.03	.93	.86	.94	.83	.60
19. I never do any reviewing except before examinations.	.09	.23	.07	.30	.36	.30	.31	.40
20. I get practice in reading aloud only when I am called on to read in class.	−.16	−.20	−.04	.23	.43	.36	.41	.70
21. I usually get through my German assignment in time that I have set aside for studying it.	.19	.17	.11	.53	.64	.49	.38	.40

Table 1.—*Continued*

	Correlation of item with:			Number of students in each grade classification				
				A	B	C	D	E
				30	42	47	29	10
	German Achievement	Psychological Examination	German Achievement minus Psychological Examination	Per cent of students in each grade classification who marked item (+)				
				A	B	C	D	E
	1	2	3	4	5	6	7	8
22. I spend most of my time studying the grammar part of each lesson.	−.22	−.28	−.06	.53	.26	.51	.48	.60
23. I am not able to apply correctly in written work the rules of grammar I know.	−.50	−.31	−.41	.10	.21	.32	.59	.80
24. I get the main facts or principles of grammar contained in each lesson and don't worry too much about details.	−.14	−.08	−.12	.63	.69	.60	.52	.60
25. I spend most of my study-time on the reading, questions, and exercises of each lesson.	.31	.39	.09	.83	.81	.51	.52	.70
26. I try to pick out examples of grammar rules as they occur in the reading texts.	.00	−.16	.00	.47	.57	.45	.48	.60
27. I think of grammar as an end in itself, not as a means to an end.	−.35	−.40	−.14	.10	.29	.30	.45	.50
28. I memorize examples along with rules of grammar.	.00	−.41	.00	.33	.33	.21	.38	.20
29. I study German the same way I studied a foreign language in high school.	.13	.20	.02	.20	.26	.28	.24	.50
30. Before I start on an exercise I look over it to see in what it is supposed to give me practice.	.00	−.36	.00	.90	.67	.75	.79	.90

31. I study declensions by writing them out until I feel that I know them.	−.24	−.39	.00	.47	.33	.45	.52	.60
32. I study declensions by repeating them to myself with my book closed.	.00	.13	.00	.67	.74	.64	.72	.80
33. I make no special effort to learn declensions, I just hope they will stick.	.24	.28	.09	.00	.12	.11	.07	.10
34. I study declensions through applying them in exercises in the text-book.	.22	−.11	.38	.77	.60	.57	.52	.40
35. I study vocabulary by writing out the new words.	.07	−.18	.23	.30	.43	.28	.52	.20
36. I study vocabulary by attempting to recall the English equivalents of German words or the German equivalents of English words from vocabulary lists.	.13	.31	.08	.67	.50	.47	.52	.60
37. I learn vocabulary by studying it in the context of reading, questions, and exercises.	−.06	−.19	−.07	.83	.78	.83	.79	.90
38. I look up unfamiliar words in a new story only after I have tried to derive their meaning from context.	−.17	−.15	−.10	.83	.79	.74	.75	.80
39. When I run into a difficult sentence or clause in new reading material I first look for the subject, verb, and object and then try to group the remaining words about them.	.54	.24	.52	.90	.83	.68	.62	.80
40. When I read a reading text for the second or third time, I begin to read ideas rather than single words.	.33	−.21	.61	.97	.88	.87	.83	.90
41. I read and re-read each reading passage until I am able to follow its idea or content *in German*.	.34	−.08	.50	.73	.52	.49	.41	.60
42. I study corrections on written work until I understand the cause of the errors and can correct them by myself.	.47	.20	.45	.96	.83	.72	.55	.60

In validating the items with respect to the scores earned on the German Achievement test, the data from ninety-three students were employed. This group comprised every student for whom both Psychological Examination and German Achievement scores were available.

When the items were evaluated on the basis of marks earned in classroom study, correlation methods were less feasible and a statistically simple procedure was employed. For each item the percents of A, B, C, D, and E students who responded 'yes' were determined. The group whose questionnaire responses were studied on the basis of grades earned for the first semester comprised one hundred and fifty-eight students, and included the ninety-three students for whom achievement and intelligence test scores were available. The number of students who made A, B, C, D, and E is given in Table 1. It should be noted that there are but ten students in the E group. Observations based upon this small group are unreliable.

In Table 1, at the right of each item and in the order indicated, may be found the following data:

1. Biserial correlation of responses on each item with scores on the Cooperative German test,
2. Biserial correlation of responses on each item with scores on the Psychological Examination,
3. Correlation of each item with the German test after partialling out the effect of the Psychological Examination,
4. The per cent of the A students in the first semester elementary German course whose response was 'yes,'
5. The percent of the B students in the first semester whose response was 'yes,'

6, 7, and 8. The respective per cents of the C, D, and E students whose responses were 'yes' to the item in question.

Since the correlation of the German test with the test for intelligence is .621, it seemed reasonable to suppose that some of the items could be correlated with the achievement criterion by virtue of a high degree of correlation between that item and intelligence. It was considered desirable to know the correlation between each item and the test for intelligence.

The question of the desirability of attempting to partial out or to control the effect of intelligence may be raised. If one were chiefly interested in the predictive value of study habit items, the value of attempting to control the factor of intelligence might not justify the effort involved. If, however, one wished to know how conducive to achievement the practice of the study procedures described by an item might

be, it would be desirable to investigate the possibility that the relation of the item to achievement might be due to the fact that the practice is characteristic of persons who are intelligent. Items which correlate with achievement under such conditions may reasonably be considered to be of less value than items which describe a performance which contributes to achievement regardless of the intelligence of the person involved.

Discussion

In our particular discussion, the term "intelligence" may be defined as equal to the total score on the Psychological Examination and the term "achievement" is defined as total score on the Cooperative German Achievement Test.[2]

A discussion of Table 1 may be made by proceeding through it item by item:

1. Responses on this item are without apparent significance.
2. The tendency to worry about the effectiveness of his study while studying chiefly characterizes the student who is making poor grades. It does not appear to be related to achievement or intelligence as measured in this study.
3. This evidence of anxiety or distraction, like that of item 2, seems to be characteristic of students who are making poor grades, and not to be related with the measures for achievement or intelligence.
4. Students who are making good grades and those who are achieving tend to feel more handicapped by distractions than others. This tendency shows a slight inverse relation with intelligence.
5. Procrastination is independent of intelligence and without effect on achievement. It may be unrelated to grades.
6. The presence of examination "jitters" is inversely related to achievement, intelligence, and scholarship. Perhaps the basic factor in this relationship is ability. Students who are not able fail to do well in courses, do poorly on exams, and as a consequence fear the exam and attribute much of their difficulty to a special kind of examination handicap. Striking exceptions to this generalization are frequently met, and it would be dangerous to overlook the possibility of "examination trouble" in students of high ability.
7. This item offers further evidence that the student who is not making good marks tends to be anxious while studying. It appears to measure much the same thing as items 2 and 3.
8. There is no apparent tendency for responses on this item to vary systematically with grades, achievement, or intelligence.
9. Although there is an inverse relation between intelligence and an affirmative response to this question, it is also apparent that the better students are inclined to respond in the affirmative to this question. This is a reasonable finding, however, and knowing what to memorize has an obvious relation with grades earned in courses, although it need not have any relationship with the scores earned on the Achievement test.
10. There is a positive relationship between grades earned and understanding what the instructor wishes to achieve. This confidence in or understanding of the instructor appears to be positively related to achievement, but it is

[2] Because of their objectivity, standard achievement tests may be superior to grades as indices of true achievement. Often, however, they possess the disadvantage of not being specific to the content of the course.

inversely related to intelligence. This may be explainable by a certain slight facetiousness which is often characteristic of the more intelligent students. It may be explainable by the fact that ordinary classroom procedure often does appear to be without aim to the very superior student.

11. Drilling material until it becomes automatic is inversely related with intelligence, and appears to contribute little to achievement or to grades; the A student is a possible exception.

12. It is interesting to note that the persons most likely to view each lesson as a step toward mastery are the very good and the very poor students. Each of these groups may be alike in the respect that they have well-defined aspirations.

13. It is apparent that the attitude expressed in this item is somewhat conducive to achievement and is correlated positively with intelligence. Most students give an affirmative response to this item. The same remarks that apply to number 12 may be made about this item.

14. Whether a student regards his assignment as a tedious task or not seems to be unrelated to achievement or intelligence. This attitude, however, seems to be most characteristic of the less efficient students.

15. Attempting to master each lesson or assignment is quite highly correlated with achievement and intelligence. It is also associated with good grades. It is interesting to note that the small group of students who failed also show a relatively high tendency to say 'yes' to this item. It is likely that mastery to the group that fails and mastery to the group that makes high grades are two different concepts.

16. The allotment of time as described in this item appears to be inversely related with intelligence, and it is slightly more characteristic of the students who do not make good grades than it is of the students who make good grades.

17. This type of preparation is correlated with intelligence. It is also characteristic to some degree of those who achieve highly and those who make good marks. Relatively few of the students say that they study in this manner.

18. Studying before the next class meeting is not particularly characteristic of intelligent students, nor does it seem to be conducive to achievement, but it does bear a relationship to good grades.

19. Reviewing only before examinations is a practice which does not appear to be related with achievement or grades. There is a slight tendency for reviewing just before examinations to be associated with higher intelligence.

20. The failing students rarely read aloud outside of class.

21. Finishing up the German assignment in the time set aside for it seems to be slightly more characteristic of the students who make good grades than the students who do not make good grades.

22. There is a very slight tendency for positive responses on this item to be negatively correlated with achievement and with intelligence.

23. Inability to apply grammar rules appears to characterize students of low achievement and who make low grades; it is also inversely related to intelligence.

24. This item is of no apparent significance.

25. This item suggests that emphasis on reading and exercises at the expense of grammar is characteristic of highly achieving students, those of relatively superior intelligence, and those who make good grades. (Refer to item 22.)

26. This item is without apparent significance.

27. Thinking of grammar as an end in itself is not characteristic of good students. It is inversely related to achievement and intelligence.

28. Memorizing examples along with the rules of grammar is without significance except that it is inversely related to intelligence.

29. Almost no one seems to study German the same way he studied his high school language with the exception of those students who failed.

30. Almost everyone tends to look over the exercise before he starts working on it, but this seems to be especially characteristic of the very good and the very poor students.

31. Writing and re-writing declensions does not seem to contribute to achievement. It appears to be inversely related to intelligence, and it appears to be most characteristic of those students who made poor grades.

32. Studying declensions by repeating them to oneself seems to be most characteristic of those students who failed. Most of the students, however, employ this method.

33. This item is without significance. There is a tendency for a positive response to be associated to some degree with intelligence and with achievement. Few students say 'yes' to this item.

34. A practice of studying declensions by applying them in the exercises is positively related with good grades.

35. Learning vocabulary by writing the new words is unrelated to the criteria we have employed.

36. The use of vocabulary lists tends to characterize the students who are unusually good or unusually poor.

37. Response on this item is of no significance; most students give an affirmative response.

38. This practice characterizes most students, and the item has no discriminatory power.

39. This systematic approach in reading is highly related to achievement, tends to be correlated with intelligence, and is also positively related to grades. There is an apparent tendency for failing students to employ this method.

40. This practice appears to be conducive to achievement, but it is not positively related to intelligence. The A students as a group give an affirmative response to this item.

41. Reading until the thought comes in German is conducive to achievement and characteristic of outstandingly good students.

42. Studying one's own work and understanding the corrections characterizes students who achieve well and also characterizes those who make good grades. There is a suggestion that the tendency is positively related to intelligence.

In this discussion, the interpretive comments are most applicable to the present population and procedure. The degree to which these findings may be applied to other situations is necessarily a function of the similarity of other situations to the present one. It is likely that the relationships between the criteria and the items would be modified in a group of students who had studied German for several years. The findings would probably be modified in a more mature group, such as graduate students.

A very important factor in determining the applicability of these findings to other schools is the similarity of the organization of the first course in German at the University of Illinois to the organization of this course in other schools. At the University of Illinois Part I of the textbook *Ich Lerne Deutsch*[3] is employed in the first semester German

[3] Ross, F., Ross, A. E., and Aron, A. W. *Iche lerne Deutsch.* New York: Harper and Bros., 1935, pp. 466.

course. "The teaching procedure based on this text may be described as a modified reading method. The primary goal of the course is to develop the student's ability to read with ease and pleasure."[4]

Conclusions

From an examination of the data in Table 1, a number of observations, many unanticipated, may be made. In examining this material it must be remembered that these data are questionnaire data and are subject to all the disadvantages inherent in data collected by the questionnaire method.

1. Items which describe or involve routine, such as numbers 11, 16, 28, 31, 32, and 36, are found to be of relatively little importance in this study.

2. Items, 1, 16, 17, 18, 19, and 21 describe the use of time. Responses on these items bear little relation to academic achievements. These findings suggest that the counselor's frequent emphasis on the exact use of time may in many instances be unwarranted.

3. There appears to be a significant relation between a student's scholastic standing and his feelings of anxiety toward the course. This may be interpreted as indicating the desirability of building up the confidence and reducing the anxiety of ineffective students.

4. It is apparent that it is desirable for the student to be fully aware of the aims of the course, and to know what relative emphasis to place on the various components of the course.

5. Although the number of significant items which are specifically descriptive of study procedures is limited, they are of such a nature as to be useful. They are items 4, 15, 23, 25, 27, 34, 39, 40, 41, and 42. One of these deals with distraction, another with attitude, another with specific difficulties, and two are concerned with the emphasis to be placed upon the various phases of the work. Five additional items deal with concrete study procedures specific to the course.

Received September 7, 1943.

[4] Larsen, Wittenborn, and Giesecke. *Op. cit.*

The Use of the Free Association Technique in the Investigation of the Stimulus Value of Trade Names

John P. Foley, Jr.

The Psychological Corporation, New York City

The free association technique, originally developed as a diagnostic method of revealing the idiosyncracy *vs.* the commonality of an individual's verbal associations, has been adapted for use in several fields of applied psychology, such as the measurement of associative reaction time, the classification of associations into various types, and the analysis of the subject's associations in order to reveal or diagnose interests, attitudes, "guilty" or technical knowledge, aptitudes, intelligence, emotional "complexes," personality traits, pathological processes, and the like, as well as in psychotherapy. The present study represents an attempt to use the free association technique in the investigation of the stimulus value of trade names,—a problem of considerable importance for the psychology of advertising, marketing, and related disciplines.

Subjects

The subjects employed in the present study were 300 George Washington University students, ranging in age from 16 to 46, with the majority falling between 18 and 21. The group was composed of approximately 40% men and 60% women. The range in educational level was from Freshman to Post-Graduate, although there was a predominance of Freshmen and Sophomores. The subjects were divided into three experimental groups of 100 subjects each, the groups being equated with respect to age, sex, and educational level.

Procedure

The procedure employed was the standard free association method, although only one stimulus word was administered to each subject and only one response word was obtained. When each subject had been provided with a pencil and blank 3″ × 5″ card, the following directions were given:

"This is a brief test of verbal association. I shall pronounce a word or name. As soon as I have pronounced it, you are to write down on your card the *first* word that occurs to you, i.e., the *first* word you think of after hearing the word or name that I pronounce. For example, if I say "window" and if you think of "door," you would write down the word "door" immediately.

"The word or name I give will be pronounced only *once;* it will not be repeated. Do not converse with your neighbor, and do not look at his card. We want your own individual reaction.

"Please remember: you are to write down the very first word you think of after hearing the stimulus word. Write only *one word.* Write it down regardless of what it is."

In one experimental group the stimulus word "Pepsi Cola" was then given orally; after which each subject wrote down his verbal response. The same procedure was followed in the other two experimental groups, except that the stimulus words "Coca Cola" and "Dixie Cola" were given, respectively. At the conclusion of the experiment, each subject was asked to indicate sex, age, and year in college.

Results

In Table 1 will be found the complete alphabetized list of associations given to each of the three stimulus names. In each case the subjects' associations have been transcribed without change, even though the response was a phrase rather than a single word. In the case of the responses "Coca Cola" and "Pepsi Cola," abbreviations and colloquialisms as well as misspellings have been classified separately under their respective names, although the total number of generic associations is also given in each case in parentheses. Since there are 100 subjects in each of the three experimental groups, each entry in the first three response columns is a percentage, and intercomparisons can be made directly. The last column represents the total number of associations made in the three experimental groups combined.

It is interesting to note at the outset that there is a surprisingly small range of associations to each of the three stimulus trade names. Only 63 different responses were made by the 300 subjects, a result which would seem to indicate that associations to such trade names are limited in number and hence readily investigated by the free association technique. The trade name "Coca Cola" seems to be associated with a slightly greater number of different things than the other two trade names, since there are 29 different responses to "Coca Cola," 27 to "Dixie Cola," and 22 to "Pepsi Cola."

Let us next consider the extent to which a given stimulus trade name is associated with the trade name of another Cola product as a response. This is the most important part of the study as far as trade name confusion and infringement are concerned. If we consider the generic totals in making such a comparison, it is found that surprisingly large percentages of the total number of associations are of this type. To the stimulus name "Pepsi Cola," 22% of the subjects gave the response "Coca Cola" (or one of its generic equivalents); there are no other "Cola"

Table 1

Response Word	Stimulus Word			
	Pepsi Cola	Coca Cola	Dixie Cola	Total
Awful		1		1
Bells	1			1
Beverage			1	1
Big	1			1
Bottle	11	22		33
Brown		1		1
Caffine		1		1
Candy	1			1
Canteen	1			1
Clear Cola			1	1
Coca Cola (generic)				
Coc		1		1
Coca			1	1
Coca Cola	16		26	42
Coke	6		2	8
(Total generic)	(22)	(1)	(29)	(52)
Cola			1	1
Cold		2	1	3
Commercial			1	1
Dixie-cup			1	1
Drink	40	32	11	83
Drinking		1		1
Drug store		2		2
Excitement	1			1
Flying	1			1
Fountain		2		2
Ginger-ale		1		1
Glass		3		3
Good	1			1
Hits	4			4
Hits the spot	3			3
Ice		1		1
Ice cream			5	5
Large bottle	1			1
Machine		1		1
Magnolia tree			1	1
Mint		1		1
Music—swing			1	1
Nickel (nickle)	3	1		4
Paper cup			1	1
Pensa-cola			1	1
Pensicola, Florida			1	1
Pepsi Cola (generic)				
Pepi Cola			1	1
Pepsi Cola		10	27	37
Pespi Cola			1	1
(Total generic)		(10)	(29)	(39)
Peter Arno	1			1
Pop	1	2	2	5
R. C. Cola		1		1
Red			1	1
Reddish-brown		1		1
Refreshes		1		1
Refreshing		1		1
Refreshment(s)		3	1	4
Root beer		1		1
Sandwich		1		1
Sherbet			1	1
Soda	2	3		5
Soft drink	2		4	6
South			4	4
Sparkling beverage			1	1
Straw		1		1
Tastes	1			1
Thirst	1		1	2
Vanilla		1		1
Water	1	1		2

responses to this stimulus. To the stimulus name "Coca Cola," 10% of the subjects gave the response "Pepsi Cola," and 1% gave the response "R. C. Cola." To the stimulus name "Dixie Cola," 29% of the subjects gave "Coca Cola" and 29% gave "Pepsi Cola;" 1% gave "Clear Cola," the only other "Cola" response made in the present study.[1]

From these figures, several interesting facts are clearly apparent. Note that there are many more "Coca Cola" responses to other "Colas" than there are other "Cola" responses to "Coca Cola," a result which would indicate that Coca Cola is the best known Cola product and that the subjects tend to think of it when they hear the name of another Cola product to a much greater extent than they tend to think of another Cola when they hear the trade name "Coca Cola." It will also be noted that the response "Coca Cola" is more frequent to the stimulus "Dixie Cola" than to the stimulus "Pepsi Cola." Similarly, the response "Pepsi Cola" is much more frequent to the stimulus "Dixie Cola" than to the stimulus "Coca Cola." In fact the responses of "Pepsi Cola" and of "Coca Cola" to the stimulus "Dixie Cola" are of equal strengths (29%). "Dixie Cola," it is found, is not given as a response to either "Pepsi Cola" or "Coca Cola." On the basis of these results it would appear that Pepsi Cola is a stronger competitor of Coca Cola than is Dixie Cola, since it is sometimes given as a response to the stimulus "Coca Cola" (10%) and at the same time "Coca Cola" is not as frequently given as a response to it (22%) as to "Dixie Cola" (29%).[2] The fact that "Dixie Cola" is not given as a response at all would tend to strengthen this conclusion.

Let us now consider the remainder of the associations, i.e., those which do not involve specific Cola trade names as responses. The most common of these responses is "Drink," this response being given in 40%, 32%, and 11% of the cases in the "Pepsi Cola," "Coca Cola," and "Dixie Cola" groups, respectively. The relatively smaller incidence of this association in the last-named group is tied up with the fact that 58% of the responses of this ("Dixie Cola") group were to the trade names "Coca Cola" and "Pepsi Cola." The only other non-trade name response which appears with great frequency is "Bottle," being given in 11% of the cases in the "Pepsi Cola" stimulus group and 22% of the cases in the "Coca Cola" stimulus group; it is interesting to note that this response does not occur at all in the "Dixie Cola" stimulus group. The greater incidence of this response to the stimulus word "Coca Cola" undoubtedly reflects the greater familiarity and perceptual distinctiveness of the Coca Cola bottle.

[1] The fact that no other "Cola" responses occurred is of interest in the light of the extremely large number of differently named "Cola" products.

[2] It would be interesting to study the effect of advertising in setting up a differentiation between two such products on the part of the consumer public.

The remaining non-trade name associations occur with small and scattered frequencies, and often fall into certain general categories. Noteworthy are those involving *color* ("Brown," "Red," "Reddish brown"), *generic terms* ("Beverage," "Pop," "Soda," "Soft drink," "Sparkling beverage"), *size* ("Big," "Large bottle"), *associated contiguous objects* ("Candy," "Canteen," "Dixie Cup," "Drug store," "Fountain," "Ginger ale," "Glass," "Ice," "Ice cream," "Machine," "Mint," "Paper cup," "Root beer," "Sandwich," "Sherbet," "Straw"), *effects* ("Refreshes," "Refreshing"), *evaluation* ("Awful," "Good"), and sound association ("Pensa-Cola," "Pensicola Florida"—both given to "Pepsi Cola"). *Geographical associations* are involved in the 4 responses of "South" and single response of "Magnolia" to the stimulus "Dixie Cola." The *influence of radio and other forms of advertising* is also indicated by the responses "Big," "Canteen," "Hits," "Hits the spot," "Large bottle," "Nickel," and "Peter Arno" to the trade name "Pepsi Cola," and by the responses "Refreshes" and "Refreshing" to the stimulus name "Coca Cola." In this latter connection it is of interest to note that the response "Thirst" is given by only 2 subjects and does not occur to the stimulus "Coca Cola," in spite of the advertising slogan, "Thirst knows no season."

Received September 20, 1943.

Job Specifications in Applied Psychology

Note: The editor will be glad to receive job descriptions and specifications in the fields of applied psychology. These will be published from time to time to provide general information for students, applied psychologists, employers, and employment agencies.

The following requisition was recently received by a number of employment agencies in the Midwest.

Corporation Title: Employee Analysis Assistant. USES Title and Code: Psychologist, 0–39.17.

Summary of Duties: Assist in the development of employee selection and placement tests including securing of criterion data, making item analyses, determination of validity and reliability, standardization, writing manuals of instructions, etc. Administer, score, and interpret tests. Some counseling and guidance. Assist in developing training programs, preparation of conference materials, etc. Assist in the development of rating scales, and make analyses of results similar to those of employment tests.

Requirements: Minimum of BA in psychology and preferably a MA in psychology, courses to include elementary and advanced statistics, tests and measurements, clinical psychology, experimental psychology, and applied, industrial, or personnel psychology. Woman preferred. Must present proof of American citizenship. Must take a physical examination.

Salary: Open, depending on training, experience, and other qualifications of applicant.

Hours of Work: Normal work week of 39 hours; 8:00 to 4:00 Monday through Friday and from 8:00 to 12:00 on Saturdays.

General Information: Opening in the Industrial Relations Department, Main Office Building, Blank-Blank Corporation. Immediate supervision will be given by industrial psychologist. Nature of work, amount and type of supervision will depend on abilities shown on the job. Some of the work will be quite routine to start, and applicant must indicate high accuracy and reasonable speed in routine work.

Desirable applicants should have good judgment, poise, pleasing appearance, desirable mannerisms, ability to gain ready acceptance and enthusiastic support of other persons, ability to express self effectively in written and oral reports. Ability to type and experience with Monroe calculating machine desirable. Applicants should give in detail all

personal data and information as to education, training, experience, interests and other pertinent evidence showing qualifications for work described.

The following requisition from a large manufacturing company in Ohio was recently received by a number of colleges and universities.

Title: Personnel Research Assistant. USES Title and Code: Psychologist, 0–39.17.

Summary of Duties: Developing and carrying out, with clerical help, statistical procedures used in psychological test construction. The greater part of the research work is designed to improve selection methods. It involves test construction, rating methods, and various other phases of personnel research.

Requirements: Training or experience in the following: elementary statistics including correlation, ability to operate calculating machines, ability to set up economical procedures for statistical calculations, courses in mathematics through calculus, courses in either education or psychology especially with respect to test construction and validation. Woman preferred. All of these requirements are desirable but not essential.

Hours of Work: Normal work week of 35 hours; 8:30 to 4:30 Monday through Friday.

General Information: Opening is in the Personnel Research Department located in the company's downtown office. Applicants should give information regarding age, marital status, salary requirements, course work in mathematics, statistics, education, and psychology, and experience in statistical computation.

Book Reviews

Jones, Harold E. *Development in adolescence.* New York, D. Appleton-Century Co., 1943, pp. xvii + 166. $2.00.

This book is a study of development during preadolescence and adolescence. Specifically, it deals with the history of one boy, John Sanders, chosen from the 200 boys and girls of the California Adolescent Growth Study. We first see John at the age of 10 and leave him at the age of 18 when he is presumably preparing to enter college. The ten chapters in the book cover John's early life in the home, describe his parents, and deal at length with John's physical, mental and social development. In the last two chapters we are given a picture of John as he sees himself, and the author's summary of John as a person.

The techniques employed in the study are numerous and well chosen, and are certainly worth the attention of anyone interested in the psychology of individual differences. A wealth of measurement data is presented, ranging from precise physical, mental and educational measurements to ratings and observations of John's social behavior. Whenever possible John's development is appraised against the norms of his group. Scores are uniformly in terms of standard deviation units, the norm for the group being set at 50 and the standard deviation at 10, so that the range of scores is approximately from 20 to 80. The use of standard scores is especially interesting since, at the present time, many social psychologists who emphasize the need for obtaining a picture of the "whole personality" decry mental measurement as such, the projective techniques being championed as the only methods in which an individual is described as a person. It is noteworthy, however, that unless John is compared with a norm it is impossible to tell whether he deviates sufficiently from the behavior of boys of his age group to be characterized as atypical.

The author does not hesitate to make use of any techniques, subjective or objective, which will throw light upon John's development. Throughout the book the various measurements are skillfully interwoven and compared. The understanding of John's backward social development at age 15, for example, is greatly enhanced when we know that he was undersized and underdeveloped physically at the same period. Chapters five and six, which present a detailed and graphic picture of John's physical development and of the growth of his mental and motor abilities, are particularly valuable. Changes in John's interests and attitudes pre-

sented in later chapters are tied in with earlier determinations by frequent cross references. One gets the impression that John's social and emotional behavior was considerably restricted by various factors, environmental as well as personal. But these restrictive factors were probably not more numerous than those faced by most adolescent youngsters.

To the reviewer, the objective and observational measures come out better than the projective techniques. The account given of John by the Rorschach Test (the records were analyzed "blind") are not too helpful; and this is true also of the description of John's personality obtained from the Murray Pictures. As judged by objective records, the conclusions drawn from these accounts are often wide of the mark, and sometimes are reminiscent of the "fortune" which one gets from the penny weighing-machines. For example, the Rorschach interpretation describes John as follows: "His superior intelligence and his strong inner life (inner creativeness) serve to stabilize his labile affectivity." When we find that repeated measures of John's intelligence give him the relatively modest I.Q. of approximately 110, it is hard to see how his intelligence can be described as "superior." Just what is meant by saying that "inner creativeness serves to stabilize labile affectivity" is beyond the ken of the reviewer.

The final picture of John, given in Chapter ten, is that of a young man, who, despite ups and downs in his social, emotional and physical development, has become a relatively effective young person. We are left with the feeling that John has somehow come through his adolescent stresses and strains successfully. The author writes: "His (John's) college years also brought a successful record in courses and an enterprising variety of outside activities. So marked an up-turn in John's personal fortunes is evidence not only of the toughness of the human organism but also of the slow, complex ways in which nature and culture may come into adaptation." The reviewer suspects that this statement represents more often than not the "pilgrim's progress" of most adolescents.

Henry E. Garrett

Columbia University

Wolff, Werner. *The expression of personality.* New York, Harper, 1943, p. xiv + 334. $3.50.

This book is primarily a report of the experimental depth studies carried out by its author since 1925. Part I introduces the experimental methods. Part II reports studies in self-confrontation concerned chiefly with the individual subject's reactions to his unrecognized forms of expression, i.e., recordings of his voice, photographs of head and hands, cinema records of gait and movement, and handwriting. Part III re-

ports studies of the influence of emotional factors on memory and presents some experimental diagnoses of personality based on expressive behavior (voice, gait, profile, hands, narrative style). Other techniques of diagnosis are given in the author's discussion of his "crossroads" tests, through which the subject's reactions (recall, perception, and decision) to a series of stories are studied. The consistent purpose of the whole program of research is to combine psychoanalysis with experimentation and to produce eventually an experimental depth psychology. In the author's words, "Our technique of self-confrontation as well as the experiments on memory are means of obtaining a subject's complex words. These words are the highway to the depth of personality. The aim of experimental depth psychology is to find methods for determining these complex words experimentally."

Readers of this book will probably agree that one of its most unique contributions to the psychology of personality is the series of studies on self-confrontation. The author finds: 1. That individual subjects frequently fail to recognize their own forms of expression. 2. Despite nonrecognition subjects tend to react distinctively to their own forms of expression. 3. The subject's reaction to his own form of expression tends to be more extreme, more extensive, and more penetrating, than his reactions to the expressive forms of other persons. The author interprets these results as a demonstration of the "individual's dynamic attitude toward himself." Huntley has provided independent experimental verification for some of Wolff's results, but places a somewhat different interpretation upon them. Positive trends appear also in the studies of memory and in other parts of these programs of research. In these, as well as in the experiments on self-confrontation, small numbers of subjects are usually employed, and many of the purely factual outcomes will require further verification.

The reader's evaluation of the book as a contribution in relation to the problems of practical diagnosis and prediction will depend to a considerable extent upon his reactions to Chapter XIV, "Experimental Diagnosis of Personality" and Chapter XV, "Indirect Verbal Methods of Assessing Intra-Personal Tendencies." From an experimental point of view, the first of these chapters seems weak. The demonstration is based upon four selected cases. The assumption that consistency in the interpretation of different forms of expression is proof of validity is interesting, but, from the point of view of your reviewer, debatable. The postulating of "two souls" or battles between "two psychic principles" to explain inconsistency as between different forms of expression, may well cause the reader to wonder how far the type of diagnosis employed here escapes from the more subjective features of psychoanalytical

interpretation. A more rigorous criterion of validity is needed, as well as a more extensive demonstration. The "crossroads" tests presented in Chapter XV are suggestive and interesting, but difficult to evaluate on the basis of the evidence as given.

In the concluding section of the book the author writes, "our procedures of uniting the point of view of depth psychology with the experimental method led us to findings which may be of service for the diagnosis of personality. The tests may give an experimental base for the development of skill in characterological ability and for the confirmation of statements about personality by experiment." Certainly the author's evaluation is conservative enough. Undoubtedly many will agree with the reviewer that, considered as a report of an extensive exploration in the field of personality in which many original and ingenious methods were employed, this is an important book.

Howard Gilkinson

University of Minnesota

Burroughs, Harry E. *Boys in men's shoes.* New York, Macmillan, 1944, p. xx + 370. $3.50.

Although written to have a wide and popular appeal, this book will be of particular interest to social workers, counselors, clinicians and guidance persons generally. It is the story of the origins, development, present program, and future hopes for the Burroughs Newsboys Foundation of Boston, a unique private institution in child guidance. In this collection of anecdotes, "case histories," comments, personal observations and philosophy, the very lack of organization lends a certain charm. Writing in a straightforward, conversational style, the author succeeds quite as well in giving an intimate picture of himself as of his Foundation. The reviewer is tempted to add that the author *is* the Foundation, so closely have Mr. Burroughs' own childhood experiences and his adult philosophy guided the conception of the program, so inextricably is his personality enmeshed in its functioning.

The book presents an intriguing conglomeration of viewpoints and ideas. The "harder-headed" will at once refer to the author's approach as sentimental and his belief in success through effort as "Horatio Alger stuff." Political and economic reformers may decry the practice of encouraging boys to respect and emulate wealthy, "self-made" men, and conservatives may disparage the intense faith in the wisdom and abilities of the street gamin. The progressive educator will find more than a hint of "formal discipline," and the essentialist will be disturbed at the author's small regard for conventional school practices. The enthusiast for psychological tests in guidance will find the book quite foreign; the paternalistic counselor will probably receive it with considerable en-

thusiasm. Scholars may deplore the absence of a professional terminology in the frequent discussions of psychological and sociological problems. Yet from the mass of seemingly contradictory opinions and practices a rather simple and ancient philosophy appears—one has only to find the great deal of good which lies within a person—every person—to do a great deal with him. Mr. Burroughs emphasizes that a wide variety of experiences with public leaders of strong personality, music and other arts, crafts, nature and the soil, will, in most boys, start the process of self-realization or regeneration, as the case may be.

The author's philosophy and techniques of guidance will be of most interest to readers of this review. In a closing chapter, Mr. Burroughs states that his entire work is founded on two principles: (1) the proper satisfaction of the ego as the touchstone to securing socially acceptable conduct, and (2) the nourishment of the individual with "certain spiritual substances . . . just as essential for the development of the soul as vitamins for the development of the body," which include realization of hopes, affection, security, exhilaration, contentment and many others. The author does not, however, discuss systematically his methods of guidance; indeed he states that "the ability to guide cannot be acquired through teaching; it must be born in a man. Education and experience help, but only if the original gift is there" (p. 101). Sheldon Glueck, who believes that many persons with special gifts for guidance can nevertheless profit from instruction in techniques, presents in his introduction to the volume a very able summary of a number of Mr. Burroughs' procedures. The reader is referred to this introduction and to the final chapter, but he must also read the body of the book if he is fully to appreciate Mr. Burroughs' own rare genius for handling boys. Again and again, in the narratives or in the stenographic reports of case conferences and interviews, Mr. Burroughs makes accurate diagnoses and useful prognoses, based apparently on great personal skill at empathizing juvenile problems and attitudes.

With no thought of presenting a complete list the reviewer has abstracted a series of principles and procedures, recognized and used by Mr. Burroughs, which seem sound in terms of modern developmental psychology. Among these are:

1. In the guidance process, the counselor is recognized as changing as well as the counselee. Even though guidance skill is looked upon as a special talent, the counselor at no point is considered all-wise or all-powerful. All who contact the boy are considered to participate in the guidance process.

2. The author does not admit to a special "system" or "method" of counseling.

3. The nature of the street trades presumably attracts aggressive, ambitious boys, hence the author's child groups are selected at the beginning in terms of certain positive qualities.

4. A wide use of anecdotal records is important in evaluating the boy's progress.

5. The changing of attitudes is considered basic to the changing of behavior. Modification of circumstances of living may be necessary to changing attitudes. To achieve this end, quite complete nutritional, medical, tutoring, and parent education services are made available by the Foundation and their use popularized among the boys and their parents. A scholarship program enables academically promising boys to begin their program of higher education.

6. Wide ranges of abilities and capacities exist in any group. Various motor, mechanical, artistic and educational skills and abilities are assumed to intercorrelate only moderately.

7. A wide variety of opportunities is necessary to insure adequate exploration of each boy's possibilities.

8. Learning by participation is stressed in the acquisition of appreciations as well as of understandings and skills. Positive rather than negative rewards are used.

Probably all educators and counselors are not agreed upon the effectiveness of such practices as the following:

1. Deliberate, direct appeals to juvenile egotism and desire for approbation and applause.

2. A very extensive use of material recognitions and prizes for desirable behavior.

3. Hero worship of successful business men, politicians, and other "self-made" individuals.

4. Vocational guidance based principally upon an interest in creative work and encouragement of imagination and pride in the finished product.

5. An extensive use of stereotypes—mottoes and slogans—as motivational devices.

While this book should cause one to appreciate more fully the possibilities of personal influence in counseling, it will not likely lead to a successful school of guidance by inspiration. There are too few individuals, interested in children, who possess Mr. Burroughs' fortunate combination of ingenuity, insight, and personal experience.

The narrative, however, testifies impressively to the quality of abilities latent in the masses of the people which may be realized, to society's gain, when opportunity for development is secured.

Dale B. Harris

Institute of Child Welfare,
University of Minnesota

AMERICAN PSYCHOLOGICAL PERIODICALS

American Journal of Psychology—Ithaca, N. Y.; Cornell University. Subscription $6.50. 624 pages annually. Edited by K. M. Dallenbach, Madison Bentley, and E. G. Boring. Quarterly. General and experimental psychology. Founded 1887.

Journal of Genetic Psychology—Provincetown, Mass.; The Journal Press. Subscription $14.00 per annum (2 volumes). 1000 pages annually. Edited by Carl Murchison. Quarterly. Child behavior, animal behavior, and comparative psychology. Founded 1891.

Psychological Review—Northwestern University, Evanston, Illinois; American Psychological Association, Inc. Subscription $5.50. 540 pages annually. Edited by Herbert S. Langfeld. Bi-monthly. General psychology. Founded 1894.

Psychological Monographs—Northwestern University, Evanston, Illinois; American Psychological Association, Inc. Subscription $6.00 per volume. 500 pages. Edited by John F. Dashiell. Without fixed dates, each issue one or more researches. Founded 1895.

Psychological Bulletin—Northwestern University, Evanston, Illinois; American Psychological Association, Inc. Subscription $7.00. 665 pages annually. Edited by John E. Anderson. Monthly (10 numbers). Psychological literature. Founded 1904.

Archives of Psychology—New York, N. Y.; Columbia University. Subscription $6.00 per volume. 500 pages. Edited by R. S. Woodworth. Without fixed dates, each number a single experimental study. Founded 1906.

Journal of Abnormal and Social Psychology—Northwestern University, Evanston, Illinois; American Psychological Association, Inc. Subscription $5.00. 560 pages annually. Edited by Gordon W. Allport. Quarterly. Founded 1906.

Journal of Educational Psychology—Baltimore, Md.; Warwick & York. Subscription $6.00. 720 pages annually. Edited by J. W. Dunlap, P. M. Symonds, and H. E. Jones. Monthly except June to August. Founded 1910.

Psychoanalytic Review—New York, N. Y.; 64 West 56th St. Subscription $6.00. 500 pages annually. Edited by Smith Ely Jelliffe. Quarterly. Founded 1913.

Journal of Experimental Psychology—Northwestern University, Evanston, Illinois; American Psychological Association, Inc. Subscription $7.00 per annum. 520 pages annually. Edited by Samuel W. Fernberger. Bi-monthly. Founded 1916.

Journal of Applied Psychology—Northwestern University, Evanston, Illinois; American Psychological Association, Inc. Subscription $6.00. 480 pages annually. Edited by Donald G. Paterson. Bi-monthly. Founded 1917.

Journal of Comparative Psychology—Baltimore, Md.; Williams & Wilkins Co. Subscription $7.00 per annum (1 volume). 500 pages annually. Edited by Roy M. Dorcus. Bi-monthly. Founded 1921.

Comparative Psychology Monographs—Baltimore, Md.; Williams & Wilkins Co. Subscription $6.00 per volume. 400 pages. Edited by Roy M. Dorcus. Without fixed dates, each number a single research. Founded 1922.

Genetic Psychology Monographs—Provincetown, Mass.; The Journal Press. Subscription $7.00. 500 pages annually. Edited by Carl Murchison. Bi-monthly. Each number one complete research. Child behavior, animal behavior, and comparative psychology. Founded 1925.

Psychological Abstracts—Northwestern University, Evanston, Illinois; American Psychological Association, Inc. Subscription $7.00. 700 pages annually. Edited by Walter S. Hunter. Monthly. Abstracts of psychological literature. Founded 1927.

Journal of General Psychology—Provincetown, Mass.; The Journal Press. Subscription $14.00 per annum (2 volumes). 1000 pages annually. Edited by Carl Murchison. Quarterly. Experimental, theoretical, clinical, and historical psychology. Founded 1927.

Journal of Social Psychology—Provincetown, Mass.; The Journal Press. Subscription $7.00. 500 pages annually. Edited by John Dewey and Carl Murchison. Quarterly. Political, racial, and differential psychology. Founded 1929.

Psychoanalytic Quarterly—Albany, N. Y.; 372–374 Broadway. Subscription $6.00. 560 pages annually. Edited by Bertram D. Lewin and others. Quarterly. Founded 1932.

Character and Personality—Durham, N. C.; Duke University Press. Subscription $2.00. 360 pages annually. Edited by Karl Zener. Quarterly. Founded 1932.

Journal of Psychology—Provincetown, Mass.; The Journal Press. Subscription $14.00 per annum (2 volumes). 800–1200 pages annually. Edited by Carl Murchison. Quarterly. Founded 1936.

Psychometrika—University of Chicago, Chicago, Ill.; Psychometric Society. Subscription $10.00. 320 pages annually. Edited by L. L. Thurstone and others. Quarterly. Quantitative methods in psychology. Founded 1936.

Psychological Record—Bloomington, Ind.; Principia Press. Subscription $4.00. 500 pages annually. Edited by J. R. Kantor and C. M. Louttit. Without fixed dates, each number a single research. General psychology. Founded 1937.

Journal of Consulting Psychology—Lancaster, Penn.; Science Printing Co. Subscription $3.00. 240 pages annually. Edited by J. P. Symonds. Bi-monthly. Founded 1937.